



Large language models for competence-based HRM: A case study in the aerospace industry

Giuliana Barba^{*} , Angelo Corallo , Mariangela Lazoi , Marianna Lezzi

Università del Salento, Dipartimento di Ingegneria dell'Innovazione, Campus Ecotekne, Via per Monteroni, s.n., 73100, Lecce, Italy

ARTICLE INFO

JEL classification:

O15
Human resources
Human development
J24
Human capital
Skills
O33
Technological change: Choices and consequences
Diffusion processes
M12
Personnel management
C45
Neural networks and related topics
D83
Information and knowledge

Keywords:

Artificial intelligence
Large language models
Competence
Human resource management
Semantic similarity
Semantic enrichment

ABSTRACT

Competence-based human resource management (HRM) emphasises the identification, development, and utilization of employee competence to boost the organizational performance, particularly in high-tech sectors that demand continuous competence advancement. Advanced artificial intelligence (AI)-based solutions, such as large language models (LLMs), are transforming competence-based HRM by streamlining job position selection, predicting emerging competencies, and designing targeted training plans, thereby enhancing knowledge sharing and transfer. However, there is a significant gap in the literature regarding comprehensive LLM-based solutions that automate the association of competence with professional roles and the semantic enrichment of corporate competence taxonomies. In this study, we present two innovative solutions: the automated semantic taxonomy enrichment methodology (ASTEM) and the role-competence embedding-based (RCE) framework. In particular, we demonstrated the effectiveness of LLMs in bridging the informational gaps by generating coherent competence descriptions and creating accurate role-competence associations through a qualitative case study involving a big company operating in the aerospace, defence, and security industry. The proposed solutions aim to reduce manual effort, improve the precision of role-competence matches, and support data-driven decision-making. This enables companies to efficiently identify the suitable candidates, develop focused training programs, and maintain a competitive edge by rapidly adapting to changes in the market and technology.

Introduction

Given the rapidly evolving business environment, marked by fast-paced technological innovation and transformative changes, companies must enhance the competence of their workforce to ensure a competitive advantage (Brunello & Wruuck, 2021). 'Competence' refers to a comprehensive set of attributes that includes knowledge, skill, motivation, and behaviour (Di Luozzo et al., 2021; Gugnani et al., 2018). These attributes must co-exist to improve the performance of employees in specific professional roles (defined as a recognizable pattern of behaviour and attitude corresponding to an identity (Turner, 2001), which enables the effective execution of tasks and responsibilities

(Boyatzis, 1982; Corallo et al., 2010). Furthermore, assigning the employees in the right roles presents significant professional development opportunities, while helping companies reduce their turnover (Mathis et al., 2017) and align with the professional standards required by the labour market (Brunello & Wruuck, 2021).

Competence-based human resource management (HRM), also known as 'talent management' (Goonawardene et al., 2010; Kurek et al., 2024), focuses on identifying, developing, and leveraging the competence of employees to enhance the organizational performance, particularly high-tech sectors that require continuous competence development to ensure a competitive advantage. This approach enables companies to better achieve their business goals (Wallo et al., 2020),

^{*} Corresponding author.

E-mail address: giuliana.barba@unisalento.it (G. Barba).

<https://doi.org/10.1016/j.jik.2025.100780>

Received 8 October 2024; Accepted 5 August 2025

Available online 13 August 2025

2444-569X/© 2025 The Authors. Published by Elsevier España, S.L.U. on behalf of Journal of Innovation & Knowledge. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

enhance market competitiveness (Mathis et al., 2017), foster innovation (Ubeda et al., 2017), and implement targeted training programs (Corallo et al., 2010).

In the aerospace sector, integrating human factors with advanced technologies, such as AI-assisted systems, is essential for meeting the industry demands and driving continuous innovation (Morandini et al., 2024; Reagan, 2021). Aerospace is a particularly challenging sector, and is characterised by high costs and risks, a vast and multifaceted supply chain, and a multidisciplinary knowledge base that requires firms to function as system integrators, thereby integrating design and production (Acha et al., 2007). The production of aerospace systems, including aircraft, engines, guided missiles, and space vehicles, involves highly dynamic interactions between physical components and information exchange. It requires efficient coordination across multiple organizational levels (Brunton et al., 2021). Consequently, aerospace organizations focus extensively on the continuous assessment of design and manufacturing competencies, and strategic allocation of human resources to optimise the engineering processes (Morandini et al., 2024). Managing the technical competencies in aerospace engineering is particularly challenging due to the highly advanced technologies, innovative materials, and knowledge-intensive processes involved. The interactions between individual system elements influences others, creating an intricate network of dependencies that must be comprehensively coordinated (Cilliers & Spurrett, 1999). Aerospace products comprise multiple interdependent components; therefore, the production processes are distributed among several companies at different supply chain levels to effectively manage the costs, risks, and technological challenges (Esposito & Raffa, 2007). Consequently, aligning competencies with evolving industry demands is essential for maintaining a competitive edge and operational resilience.

Structured taxonomies are defined as a hierarchical tree-structure in which concepts are organised based on the semantic meaning (Caratozzolo et al., 2023). Organizations can leverage structured taxonomies to systematically manage competence-related information (e.g., hard and soft skills), thereby ensuring that they align with evolving competence demands and facilitate accurate role assignments (Kimble et al., 2016). This structured approach is crucial in industries where technological advancements and regulatory requirements necessitate continuous workforce adaptation (Konrad et al., 2022; Morandini et al., 2024). The aerospace sector presents a unique case for analysing the role of competencies owing to its high level of complexity and dependence on cutting-edge technology. Limited research has been conducted on the systematic management of competencies despite its importance. Bridging this gap through structured competence management frameworks enhances the operational efficiency, promotes innovation, and ensures that organizations remain at the forefront of disruptive advancements.

Advanced artificial intelligence (AI)-based models, such as large language models (LLMs), drastically transform competence-based HRM, by enhancing various aspects including recruitment, training, and resource allocation. They serve as powerful tools for accurate analyses of competence and personnel information, while improving knowledge sharing and transfer (Kurek et al., 2024; Malik et al., 2023). The adoption of LLMs in HRM, leveraging their considerable natural language processing (NLP) capabilities, supports a wide range of applications, from streamlining the process of selecting the ideal job position based on competence (Martínez & Fernández, 2019), predicting emerging job competence (Sheriff & Sevukan, 2023), to developing targeted training plans (Santana & Díaz-Fernández, 2023). This helps in establishing a more genuine and engaging work environment (Shivanjali et al., 2019).

Despite the significant technological advances in the field of competence-based HRM, limited research has been conducted on the development of automated competence-role association solutions and guidelines in the taxonomy semantic enrichment process, which is essential for maintaining relevance in rapidly evolving industries. In particular, studies conducted on text generation using LLMs (e.g.,

August et al., 2022; Nguyen et al., 2021; Velásquez-Henao et al., 2023) introduced various approaches to effectively employ prompt engineering methods. However, none of these works focused on the semantic enrichment of taxonomy, which is defined as the process of integrating new and contextually concepts, ensuring that their placement is neither overly general nor excessively specific to maintain the structural relevance of taxonomy (Arslan & Cruz, 2022).

Consequently, practical methodologies have not been developed on using pre-trained embedding to map competencies to roles by exploiting the existing competence and role taxonomies in organisations. For instance, Amin et al., 2022 employed pre-trained embedding models by exploiting an organization skill taxonomy to optimise the research request assignment process. Various other studies (e.g., Gughani & Misra, 2020; Kurek et al., 2024; Skondras et al., 2023) employed these models to improve the recommendation systems using data obtained from job postings on social media. However, research has not been conducted on identifying the ideal competence set for a job role by exploiting the latest advances in AI and actual company data based on a ground truth.

In this study, we proposed the automated semantic taxonomy enrichment methodology (ASTEM) that leverage LLMs to automate the enrichment of corporate competence taxonomies addressing the existing gap in literature. Furthermore, we established the role-competence embedding-based (RCE) framework to automate the association of roles with a set of relevant competencies by integrating two disconnected business taxonomies.

We conducted a qualitative intrinsic case study involving a company in the aerospace, defence, and security sector to test the ASTEM and RCE framework using real business taxonomies (i.e., competence and role taxonomies). This sector is characterised by complex regulatory environments, advanced technological demands, and multi-layered organizational structures, making it crucial to analyse innovative competence management solutions within their real-world setting (Corallo et al., 2012; Morandini et al., 2024). The data collection process involves semi-structured interviews with company experts and the iterative analysis of enriched competence taxonomy descriptions and role-competence associations.

Therefore, in this study, we aim to demonstrate the key concepts of LLMs, making complex AI concepts accessible to HRM professionals. Furthermore, we emphasised the significance of prompt engineering for the optimisation of LLM performance and proposed the ASTEM and RCE framework as innovative solutions to automate competence-based HRM processes. In particular, the proposed solutions aim to automate the semantic enrichment process of taxonomies and the association of roles with competence in the organisation. This automation reduces the time and resources required for manual updates, enhances the precision of role-competence matches, and supports data-driven decision-making. Thus, companies can more effectively identify the suitable candidates, develop targeted training programs, and maintain a competitive edge by rapidly adapting to the market and technological changes.

The paper is structured as follows. Section 2 'State-of-the-Art' presents a detailed description of the state-of-the-art in AI focusing on LLMs and their architecture, their impact on the competence landscape, and the previous studies conducted on the use of LLMs in HRM. Section 3 'Research methodology' outlines the research methodology of the study. Section 4 'Proposed frameworks' presents a detailed description of the ASTEM and RCE framework. Section 5 'Case study results' presents the case study, describes the implementation of the proposed solutions in the aerospace company, and the highlights challenges faced by the proposed methods. Section 6 'Discussion' highlights the research and practical contributions of this study, along with the main limitations and future trends. Lastly, Section 7 'Conclusions' presents the conclusion of the study.

State-of-the-Art

AI and competence-based HRM in industry

Technological advancements, combined with process digitisation and automation, are drastically transforming the competence landscape in the labour market, driving the development of new competencies while rendering others obsolete (Frierson et al., 2023). This phenomenon is particularly evident in high-tech industries, where rapidly changing technologies require the constant acquisition of new competencies to ensure a competitive advantage (Kurek et al., 2024). The increasing complexity and dynamism in global markets requires companies to be highly flexible to constantly adapt to the changing conditions, thereby requiring new and updated job profiles and related competencies (Frierson et al., 2023).

Based on the dynamic capabilities theory (Teece et al., 1997), organizations must develop and refine their ability to sense, seize, and reconfigure resources, including human capital, to respond effectively to technological shifts and market uncertainties. Competence-based HRM becomes a crucial mechanism in this process, as companies that effectively identify, develop, and allocate talent can better navigate industry disruptions and maintain a competitive advantage. Employee productivity primarily depends on the strategic alignment of competencies with business objectives and technological advancements, thereby reinforcing the requirement for systematic competence management (Bafna et al., 2019; Corallo et al., 2010). The competence-based approach facilitates precise role assignments, while optimising task execution, reducing HR-related costs, and increasing worker motivation and satisfaction (Gupta & Kumar, 2024). Furthermore, by reducing the turnover associated with job dissatisfaction and fostering a meritocratic work environment, firms can enhance their ability to reconfigure workforce capabilities corresponding to external changes, which aligns with the core tenets of dynamic capabilities.

Additionally, the knowledge-based view (KBV) of the firm (Grant, 1996) indicates that knowledge is the most strategically significant resource in an organization. Within this framework, competence identification and mapping become essential for firms that aim to effectively integrate, share, and apply knowledge. The ability to harness employee expertise and align it with evolving business requirements is essential for maintaining a competitive advantage (Shivanjali et al., 2019). However, limited research has been conducted on competence-based management, as workforce knowledge has not been recognised as a key determinant of organizational success in the increasingly competitive environments (Cao & Zhang, 2022).

The conventional HRM methods used to analyse text data (e.g., from interviews and questionnaires), such as topic models (e.g., latent Dirichlet allocation) or linear regression techniques, do not consider the underlying context of words in sentences (Arslan & Cruz, 2022; Lin et al., 2023). Furthermore, job role and competence descriptions established within the company are typically subjective, and can present ambiguities (Jorzik et al., 2024).

LLMs comprise an advanced category of solutions in the AI field, characterised by the usage of neural networks trained using billions of parameters on large datasets (Lin et al., 2023). These data include text from books, articles, and websites, and enable models to learn the grammar, context, and nuances of language (Liu et al., 2023).

In 2017, Vaswani et al., 2017 introduced the transformer architecture, a novel approach to NLP tasks based on the self-attention mechanism and encoder-decoder architecture framework. The self-attention mechanism enables the model to focus on different parts of the input text sequence while processing each element, thereby capturing dependencies between words based on their position in the sequence. A notable example of the progress of LLMs is the Generative Pre-Trained Transformer (GPT) model, which has demonstrated advanced capabilities in NLP and coherent text generation (OpenAI et al., 2023).

In these LLMs architectures, the encoder helps in determining and

extracting relevant information from the input text by producing a continuous representation (text embedding) (Iqbal & Qureshi, 2022), which contributes to applications in sentiment analysis, topic classification, spam detection and text generation (Kurek et al., 2024). When the architecture includes the decoder, the vector representation created by the encoder is used to generate a sequence of words, where each word depends on the previous ones due to the self-attention mechanism (Vaswani et al., 2017). However, not all LLMs adopt the encoder-decoder structure. For example, GPT uses only the decoder architecture, feeding the input data directly into the decoder and still achieving excellent results due to the self-attention mechanism (Kurek et al., 2024).

Adopting LLMs in HRM automation is a strategic application of AI that concurs with both the dynamic capabilities theory and KBV by enhancing the ability of an organization to dynamically analyse, allocate, and develop workforce competencies (Malik et al., 2023). Business applications include matching the job role requirements with the organizational requirements (Jorzik et al., 2024), thereby optimising competence-based job assignments (Martínez & Fernández, 2019), predicting emerging competencies (Sheriff & Sevukan, 2023), and developing training programs corresponding to workforce evolution (Santana & Díaz-Fernández, 2023).

Prior to recent advances in NLP, Chen and Chien (2011) developed a data-mining framework for talent identification and job assignment in the manufacturing sector. Similarly, Dickson and Nusair (2010) analysed the use of AI in HRM in the hospitality industry, demonstrating the ability of these technologies in streamlining the recruitment process and improving employee retention. Recently, Mahmoud et al. (2019) evaluated the implementation of LLM-based chatbots for social media screening into the HRM systems, thereby reducing the time spent in hiring processes and providing valuable information on the candidates. Conversely, Mathew et al. (2018) analysed the capabilities of LLMs in predicting the suitability of a candidate for replacing a qualified worker vacancy. Lastly, Gupta and Kumar (2024) developed a framework to leverage LLM-based AI solutions to overcome the bias in recruitment, analyse employee data for competence development programs, and monitor employee well-being through chatbots and wearable devices.

Pre-trained embedding models to associate job roles and competence

Although LLMs with only-decoder architecture, such as GPT, can produce text representations, they are not ideal for tasks such as text retrieval and matching (Wang et al., 2022). Conversely, pre-trained models comprising only encoder architecture produce embeddings, which are vector representations that can capture semantic nuances in sentences; they contribute significantly to various NLP tasks, such as the retrieval of large-scale information (Wang et al., 2022). These pre-trained embedding models, trained using large amounts of data and capable of performing various types of tasks, such as classification, clustering, and semantic similarity, also present excellent performance in specific tasks, such as candidate and job position matching (Kurek et al., 2024).

Several organizations already have taxonomies of competence that can be enhanced using embedding models, thereby facilitating efficient retrieval and matching with other information, such as job profile descriptions (Kurek et al., 2024) and tasks (Amin et al., 2022). For instance, Kurek et al. (2024) transformed job descriptions and candidate profiles into multidimensional embedding vectors, and calculated the scalar product between the vectors to determine the degree of similarity. A higher scalar product indicates greater similarity, indicating a stronger match between the job requirements and candidate qualifications. Similarly, Amin et al. (2022) employed the embedding model to match the research requests and skills based on a real business taxonomy.

Furthermore, Arslan and Cruz (2022) and Nikishina et al. (2022) employed embedding to enrich and populate business taxonomies using the data obtained from external sources (such as articles and lexical

taxonomies) yielding promising results. Several previous studies focused on embedding models to improve systems for recommending ideal candidates for specific job positions. For example, Lin et al. (2020) developed a human resource planning method that suggests suitable candidates for a given position. Similarly, Bafna et al. (2019) established a system to map the competence of a candidate to the corresponding job requirements and suggest specific training in case of a gap, demonstrating the effectiveness of LLMs for analysing the curricula of candidates by capturing the meaning of sentences, even with different terminologies. Furthermore, Clavié and Soulié (2023) proposed a system to extract the competence based on the job vacancies and rank the most suitable candidates. Conversely, Qin et al. (2020) developed a framework called topic-based ability-aware person-job fit neural network (TAPJFNN) for targeted talent search and job recommendation, whereas Martínez and Fernández (2019) established an intelligent system that identifies emerging talent and optimises the hiring process by excluding biases and distortions such as physical appearance, tattoos, gender, or race. Lastly, Dehbozorgi and Parizi (2023) proposed a training course recommendation system that identifies student learning gaps at the course level and recommends suitable courses to prepare them for the desired job positions.

Table 1 presents an overview of the contribution of embedding-based AI approaches to HRM. It highlights their role in enhancing taxonomies, improving candidate-job matching, and optimising the recruitment processes. Although these approaches present increased precision and efficiency, they also present challenges corresponding to the data quality, contextual understanding, and potential biases in training datasets.

LLMs for semantic enrichment process and the role of prompt engineering

LLMs are widely implemented in generative AI owing to their ability to capture complex word relationships and generate texts that reflect the semantic structure of concepts (Nguyen et al., 2021). This capability is crucial for semantic enrichment, and involves improving the understanding and representation of the meaning of texts by adding semantic information (Arslan & Cruz, 2022). Typically, manual semantic enrichment presents significant challenges in terms of the time and effort, with risks of bias or inaccuracy (Arslan & Cruz, 2022), coverage problems for specific information, and the need for frequent updates

Table 1
Contributions, strengths, and limitations of embedding-based AI approaches in HRM.

Contribution	Strength	Limitation	References
Enhancing Taxonomies with Embeddings	Organizations can facilitate more precise candidate-task matching	Requires large, well-labelled datasets and depends on the relevance of the data sources used	(Amin et al., 2022; Wang et al., 2022)
Embedding-Based Candidate Matching and Job Recommendation	Embedding models transform job descriptions and candidate profiles into vectors, enabling similarity calculations to assess candidate-job fit	The scalar product similarity approach may not fully capture the contextual meaning and human implicit knowledge	(Kurek et al., 2024)
Bias-Free Talent Identification and Recruitment Optimization	AI systems reduce hiring bias by focusing on skills and qualifications rather than demographic attributes, ensuring fairer hiring practices	Biases in training data can still affect AI-driven recruitment	(Qin et al., 2020; Ubada et al., 2017)

(August et al., 2022).

The advanced knowledge gained during the training of LLMs enables the generation of accurate and contextually relevant definitions for various terms and concepts (Eriksson & Jönsson, 2023). Large vector representations help in capturing complex relationships between words, contributing to definitions that reflect the semantic structure of concepts (Vartinen et al., 2022). Additionally, the autoregressive approach in text generation enables LLMs to construct long and complex definitions, thereby ensuring consistency and cohesion in the generated text (Eriksson & Jönsson, 2023; Lesage et al., 2024).

Recent studies demonstrated the effective reasoning ability of LLMs, particularly GPT-4, in answering medical rheumatology (Madrid-García et al., 2023) and engineering (Pursnani et al., 2023) questions without requiring specific training. However, limited research has been conducted on leveraging LLMs to simplify the semantic enrichment process. For instance, in fields such as online commerce, these models are used to generate product descriptions on websites (Nguyen et al., 2021) or to automatically identify sustainable product features in e-commerce (Roumeliotis et al., 2023). In particular, no relevant studies have been conducted in the field of competence-based HRM. Magron et al. (2024) proposed a framework to generate synthetic job postings to enhance matching with European skills, competencies, qualifications and occupations taxonomy (ESCO, 2017). Additionally, Skondras et al. (2023) employed a GPT model to generate synthetic curricula to train models to classify them into various occupational categories.

The paucity of studies conducted on the application of these models for text generation in business settings typically corresponds to the possibility of LLMs generating unwanted and inconsistent responses (Chen et al., 2023; Dahlkemper et al., 2023). However, Polverini and Gregorcic (2024) reported that the ineffectiveness of LLMs in performing tasks can be attributed to inadequate question formulation. Extensive research has been conducted on prompt engineering and its impact on the performance of LLM solutions (Chen et al., 2023; White et al., 2023), and OpenAI presents a dedicated web space to prompt engineering guidelines (OpenAI, 2023). Prompt engineering is defined as a methodological strategy for optimising the performance of LLMs through the careful design of textual prompts (Brin et al., 2023; Polverini & Gregorcic, 2024).

Extensive research has been conducted on the definition and development of this concept, demonstrating the significance of prompt design in improving the performance of LLMs (e.g., Liévin et al., 2023; Polverini & Gregorcic, 2024; Ranade et al., 2024; Velásquez-Henao et al., 2023). For instance, the chain-of-thought approach significantly improves the model capabilities via step-by-step resolution of queries and reducing the probability of generating hallucinations (i.e. fabricated or factually incorrect outputs) and illogical responses (Wei et al., 2023). Moreover, the Socratic methodology, which employs inductive, deductive, and abductive approaches to effectively improve the output quality and consistency (Chang, 2023). Furthermore, Velásquez-Henao et al. (2023) developed the goal prompt evaluation iteration (GPEI) framework, a four-step iterative methodology that can guide users to design prompts based on output evaluation. It has also been demonstrated that providing the LLM with contextual information enables it to generate more specific and relevant responses (Nguyen et al., 2021). Another key element during prompt design is the selection of pattern (i.e., output automater, persona, visualization generator, recipe, or template) that helps in optimising the quality and relevance of the generated responses (White et al., 2023). Lastly, when designing the prompt parameters, such as temperature (i.e., a parameter that controls the randomness of word selection in text generation) and top-p (i.e., a parameter that sets a probability threshold limiting word selection), that influence the probability distribution of words during text generation must be considered (Lesage et al., 2024). A lower temperature produces more concentrated and deterministic responses, whereas a higher temperature encourages more diverse and creative responses. The top-p parameter, which limits the word selection to the most likely words, enables the adjustment of

the consistency of the responses (Lesage et al., 2024; Tam et al., 2022). The combined usage of these parameters helps in optimising the model to generate sequences of sentences of arbitrary length, thereby avoiding unwanted loops and maintaining consistency with the prompt (Eriksson & Jönsson, 2023; Lesage et al., 2024).

Table 2 the main strengths and limitations of LLM-based methods in enhancing semantic structures through text generation and prompt engineering.

Research methodology

In this study, we analysed the latest advances in AI, particularly in the application of generalist LLMs without specific training, for competence-based HRM solutions. We aim to leverage the potential of LLMs for implementing semantic enrichment of competence descriptions and the association between role information and related competence from business taxonomies.

To this end, we defined and designed ASTEM (Barba et al., 2025) to guide practitioners of LLMs to effectively generate consistent outputs by integrating various Socratic prompt engineering methodologies (Chang,

2023), the GPEI framework (Velásquez-Henao et al., 2023), and OpenAI guidelines (OpenAI, 2023), along with contextual information (Nguyen et al., 2021) and prompt pattern model (White et al., 2023). Conversely, the RCE framework aims to create semantic connections between different pieces of information, specifically roles and competence, using pre-trained embedding models known for their effectiveness in semantic similarity tasks (Amin et al., 2022; Kurek et al., 2024).

We selected a qualitative investigation approach suitable for addressing the requirements of a specific context to analyse the ASTEM and RCE applications (Stake, 1995; Yin, 2017). A case study is particularly appropriate for analysing contemporary events and non-controllable units of analysis (Yin, 2017). We aim to demonstrate a real practice of particular relevance owing to the characteristics of the industry and the current research trends on HRM and LLM. Therefore, the case study is guided by the pragmatism knowledge claim that is problem-centric. Consequently, we focused on the problem and the methods used to solve it in a real organizational setting (Creswell, 2022). Multiple sources of evidence, such as interviews with managers and direct observations of researchers, were used to increase the construct validity of the case study (Yin, 2017). The case study was conducted from January to June 2024 in collaboration with a large company operating in the aerospace, defence, and security sector. Fig. 1 presents the research steps used in this study based on the guidelines provided by Crowe et al. (2011) and Kekeya (2021).

First, we defined the case study to set the research objectives and case boundaries. We primarily aim to investigate how AI can be used to enhance the alignment between professional competence and roles, as well as to semantically enrich existing corporate taxonomies. A qualitative case study involves an in-depth contextual analysis of a specific instance within its real-life setting (Crowe et al., 2011). This design was well-suited for our study because it helped in analysing the nuanced challenges and opportunities faced by a major aerospace company in implementing advanced AI-driven competence management solutions. The intrinsic nature of the case study helps in obtaining a comprehensive understanding of the subject matter (Baxter & Jack, 2015).

The second step comprises the selection of the case study (Crowe et al., 2011). The aerospace sector is characterised by rapid technological advancements and evolving competency demands. Therefore, it strongly depends on qualitative insights to evaluate the adaptability of innovative AI-based models in dynamic contexts (Morandini et al., 2024). These insights are essential for enhancing competence management and optimising the information management processes. This complexity presents an excellent testing ground for ASTEM and RCE, as advanced AI-based models can effectively capture domain-specific knowledge (Excoffier et al., 2024; Xianming & Jing, 2023).

Therefore, we selected a large company operating in the aerospace, defence, and security sector that employs over 11,000 engineers worldwide owing to its relevant contribution in terms of the interest, data availability, expertise and its proximity to the researcher's network (Creswell, 2022). By focusing on real-world documents and engaging with domain experts through interviews, the case study helped in capturing the interplay between information management through taxonomies and AI-based solutions to promote competence management. This approach also facilitated the testing and validation of ASTEM and RCE under authentic conditions. Furthermore, despite the highly technical nature of the aerospace sector, the findings of this case study enable their adaptation to broader use cases, such as corporate training, educational content evaluation, and cross-industry talent development.

The third step involves data collection (Crowe et al., 2011; Kekeya, 2021). A qualitative case study helps in capturing in-depth data using methods such as interviews, observations, and document analysis, thereby uncovering subtleties and contextual factors that may not be obtained through large-scale quantitative surveys (Stake, 1995). In particular, this step comprised a series of interviews, including three unstructured interviews to analyse the company's priorities and challenges, followed by seven semi-structured interviews to guide

Table 2
Contributions, strengths, and limitations of LLM-based text generation methods.

Contribution	Strength	Limitation	References
Generative AI and Semantic Enrichment	LLMs improve semantic enrichment by capturing complex word relationships and generating definitions that reflect the semantic structure of concepts	LLM-generated definitions may sometimes lack specificity or contextual accuracy	(Eriksson & Jönsson, 2023; Nguyen et al., 2021)
Autoregressive Text Generation	LLMs construct long and complex definitions, ensuring consistency through autoregressive generation	May still generate hallucinations or inconsistencies, requiring human oversight	(Lesage et al., 2024; White et al., 2023)
Domain-Specific Applications	LLMs automate text generation for product descriptions, competency frameworks, and job postings, enhancing efficiency in various business applications	Limited research on LLMs for HRM. Outputs in business settings may be inconsistent or irrelevant without domain-specific fine-tuning	(Madrid-García et al., 2023; Pursnani et al., 2023; Skondras et al., 2023)
Prompt Engineering for LLM Optimization	Well-crafted prompts significantly improve LLM performance, reducing hallucinations and enhancing output relevance	Poorly designed prompts can lead to misleading, verbose, or inconsistent outputs. Requires expertise to craft effective prompts	(B. Chen et al., 2023; Wei et al., 2023; White et al., 2023)
Parameter Tuning for Controlled Text Generation	Adjusting temperature and top-p enables control over creativity vs. determinism in text generation, improving coherence and relevance	Incorrect parameter settings may lead to inconsistencies and multiple attempts are needed before finding the optimal combination	(Lesage et al., 2024; Tam et al., 2022)

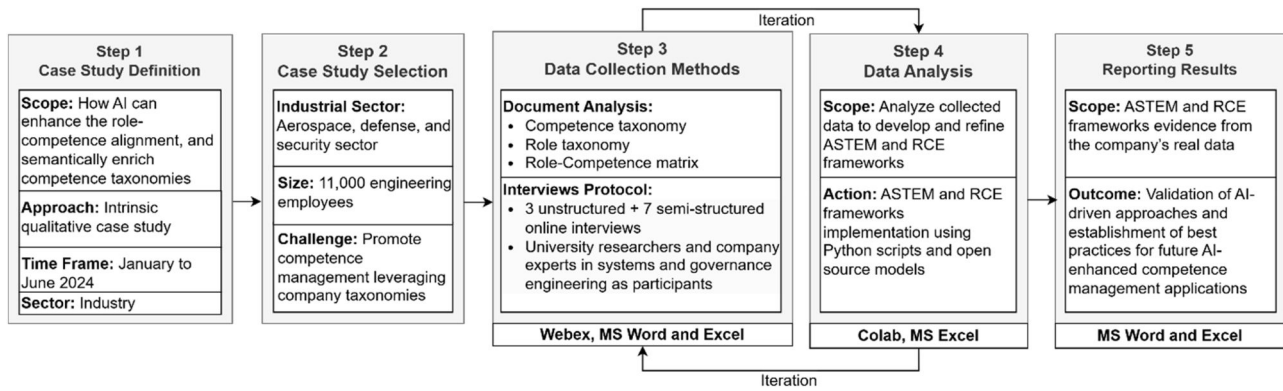


Fig. 1. Research methodology steps.

discussions towards the refinement of the ASTEM and RCE framework. These interviews were conducted online using the Webex platform by a team comprising university researchers and experts in the systems and governance engineering department from the selected company. Additionally, the company documents (i.e., competence and role taxonomies, role-competence matrix) were analysed to capture qualitative information regarding the company's perspective on competence management requirements.

The fourth step involved data analysis (Crowe et al., 2011; Kekeya, 2021). The insights gained in the previous phase contributed significantly to identifying operational strategies to enhance competence management through the effective use of advanced AI technologies. The third and fourth phases are repeated iteratively to refine the proposed solutions and implement them accurately with the real business data (Crowe et al., 2011; Kekeya, 2021). Subsequently, we developed and implemented ASTEM and RCE using Google Colab (a cloud-based platform to write and run Python code).

The last phase of the methodology involves reporting the results obtained from the case study (Crowe et al., 2011). ASTEM and RCE

provided robust evidence from the real data of the company, demonstrating the effectiveness of AI-driven approaches. The implementation of the proposed solutions validated the findings, and established the best practices for future applications in AI-enhanced competence management.

Proposed frameworks

The ASTEM and RCE frameworks are designed to work in tandem (Fig. 2), with each serving distinct but complementary functions within the competence-based HRM. In this case study, ASTEM is used to generate enriched, contextually accurate descriptions for each competence within a corporate taxonomy by leveraging the capabilities of the LLMs. Once produced, these enriched descriptions serve as the input for the RCE framework, which uses pre-trained embedding models to automatically associate these enriched competencies with the relevant job roles. The interaction between these two frameworks ensures a continuous, automated process for updating and maintaining the accuracy of the role-competence associations. We demonstrated the unique

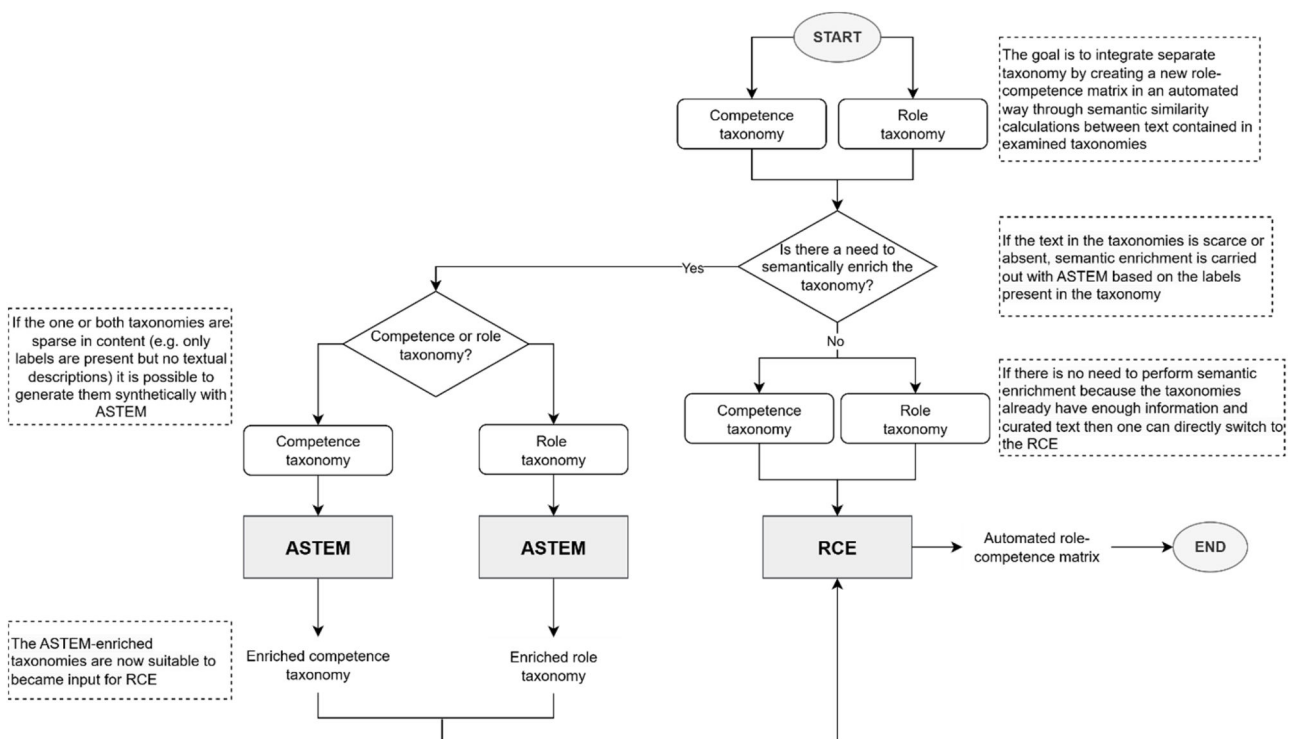


Fig. 2. ASTEM and RCE framework interaction.

methodologies and technological foundations of ASTEM and RCE in the following sections. This separation presents a clearer understanding of their individual contributions and the specialised processes involved in each, and demonstrates their combined application in the case study.

ASTEM

The ASTEM framework (Fig. 3) is an iterative methodology designed to automatically improve the semantic enrichment descriptions process of a corporate taxonomy using LLMs. The steps involved in ASTEM include: 1) preliminary taxonomy analysis, 2) prompt design (divided into system prompt and user prompt design), and 3) output evaluation. ASTEM is primarily used to reduce the manual effort involved in semantic enrichment and to improve the contextual linkage of the taxonomy elements to specific job roles and competence. This represents a significant improvement over conventional methods that employ topic modelling of retrieved information from external sources (e.g., websites, document) to add them into a taxonomy. This methodology can be applied when the taxonomy under consideration already has starting descriptions that must be enriched with more contextual information. They can also be applied when there are no descriptions. In the latter case, the performance of the model may be worse as it must generate text based on minimal information, such as the labels of the taxonomy alone.

In the preliminary taxonomy analysis phase, we defined clear objectives for the enrichment process. This phase involves a detailed assessment of the existing taxonomies to identify areas that lack detail or contain errors. Obtaining a comprehensive understanding of the current state of a taxonomy and determining the specific concepts that require semantic improvement to prepare the taxonomy is essential for effective enrichment. During this phase, control questions must be established (Table 3) based on the CRAAP test for to evaluate the information (Blakeslee, 2004). However they must also be applicable for the qualitative assessment of AI-generated content and prompt design (Newcastle University, 2024).

The prompt design phase involves designing effective queries that guide the LLM to generate accurate and relevant content. This includes selecting the most appropriate methodological approach (e.g., Chang, 2023; Velásquez-Henao et al., 2023; White et al., 2023) based on the requirements identified in the first phase. Relevant information, such as the application scenario and end goal, is added to the system prompts to enrich the context, thereby improving the ability of the model to produce useful responses (Nguyen et al., 2021). It is essential to define the required task specify how it must be performed within the system

Table 3

Control questions from CRAAP test to evaluate GenAI outputs (Blakeslee, 2004; Newcastle University, 2024).

Criteria	Description	Control Questions
Currency	Assess how recent the information is and whether it is up-to-date corresponding to the context of the examined topic	<ul style="list-style-type: none"> Does currency matter for the topic you are investigating? How up to date is the dataset used to train the AI tool?
Relevance	Examine how relevant the information is to specific needs or purpose	<ul style="list-style-type: none"> Is the information generated relevant to the task you are performing? Is the response providing the information at a level that is appropriate for your task? Is it useful because it provides a good definition and background knowledge?
Authority	Consider the competence of the author or source, e.g. their qualifications or affiliations	<ul style="list-style-type: none"> Is there evidence or references to source material that can back up the generated content? Are the sources reputable?
Accuracy	Determine the accuracy of the information, checking whether it is supported by evidence, free of errors and has an objective tone	<ul style="list-style-type: none"> Are there any factual errors or details that appear to be unsupported by evidence? Is the information reliable?
Purpose	Analyse why the information was created, considering possible bias, persuasive or commercial intentions	<ul style="list-style-type: none"> Is there any obvious bias in the response? Does the output read like fact or more of an opinion?

prompt for the LLM. In the user prompt, we included information on the hierarchical relationships (e.g., category membership and any descriptions already present) of the taxonomic element to be semantically enriched. During this phase, we set parameters such as the temperature, which generates more deterministic and less random word sequences, at low values (OpenAI, 2023).

The evaluation phase involves refining the prompts based on the qualitative analysis of a random sample of output based on the control questions established in the first phase. Additionally, a further check can be implemented by determining the semantic similarity between the enriched descriptions obtained using the LLM and each branch of the source taxonomy based on the method proposed by Excoffier et al. (2024). This helps in verifying that each enriched description was associated with the corresponding competence label via its unique code. Furthermore, if the descriptions are already present in the taxonomy, the outputs can be further verified by comparing the semantic similarity between the representative embeddings of the generated texts and the

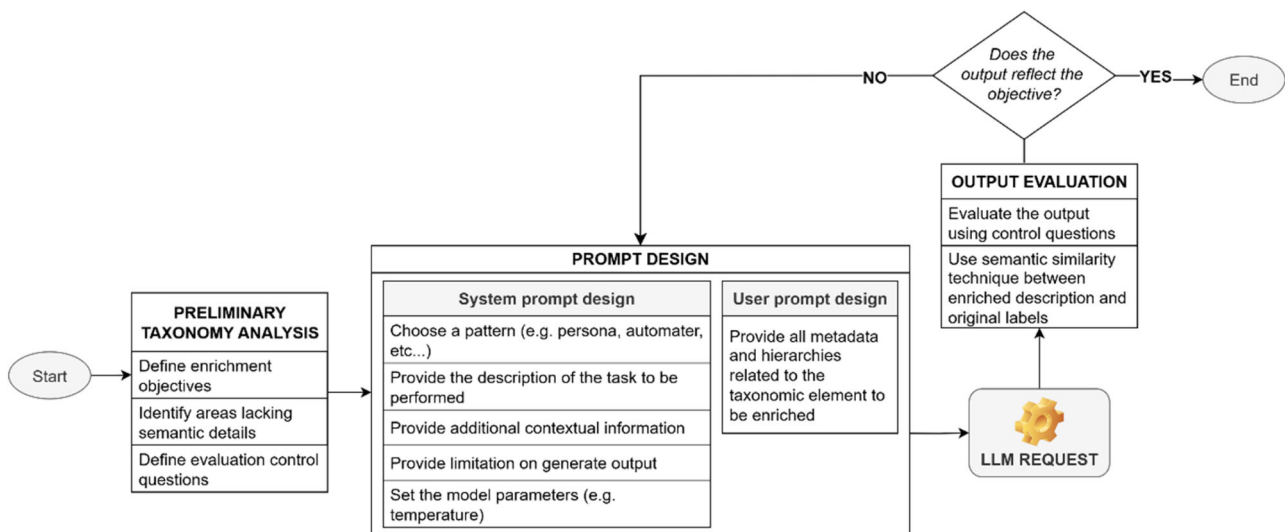


Fig. 3. ASTEM for enriching the corporate taxonomy element descriptions.

original ones to demonstrate that the generated text does not vary significantly from the original text (Celikyilmaz et al., 2021).

The prompts are iteratively modified based on the feedback received to continuously improve the quality and accuracy of the generated descriptions, thereby ensuring that the responses are contextually appropriate.

RCE

Fig. 4 depicts the RCE framework, which was developed to automatically associate each job role with a suitable and consistent set of competence. We used two reference taxonomies for this purpose: the taxonomy of professional roles and the taxonomy of competence. The steps involved in RCE include: 1) data pre-processing, 2) embedding vectors to process text corresponding to the job roles and competencies, 3) calculating the semantic similarity between these texts, which involves computing the dot product between the role and competence embedded vectors, and 4) re-ranking the results to identify the most appropriate competence sets for each role within the corresponding taxonomies.

The proposed framework involves using datasets in the Excel format that represent the taxonomies of the professional roles and competence.

In particular, the role taxonomy comprises an identifier code for each role and the associated textual strings (i.e. professional family, macro-role, activity descriptions, and mission), which are all placed in the appropriate columns. Similarly, the competence taxonomy includes an identification code for each competence, membership categories, and descriptions.

Pre-processing must be first performed to effectively apply the embedding models (Amin et al., 2022). For the role taxonomy, we concatenated all the textual strings for each role, including the professional family name, macro-role, role, role description and activities. This process is crucial as it enables the embedding model to better capture the overall semantic meaning of each role and competence, thereby increasing the probability of appropriate association. This process produces a two-column dataset comprising the unique identification role code and a concatenated description where each row in this dataset represents a specific role. The same procedure is applied to the competence taxonomy.

The embedding model must be carefully selected. There are several open source models available on <https://huggingface.co/spaces/mteb/leaderboard>; however, they face various limitations regarding the number of tokens that they can process (from 512 to 4096). Therefore, not all the text may actually be processed, presenting a significant loss of

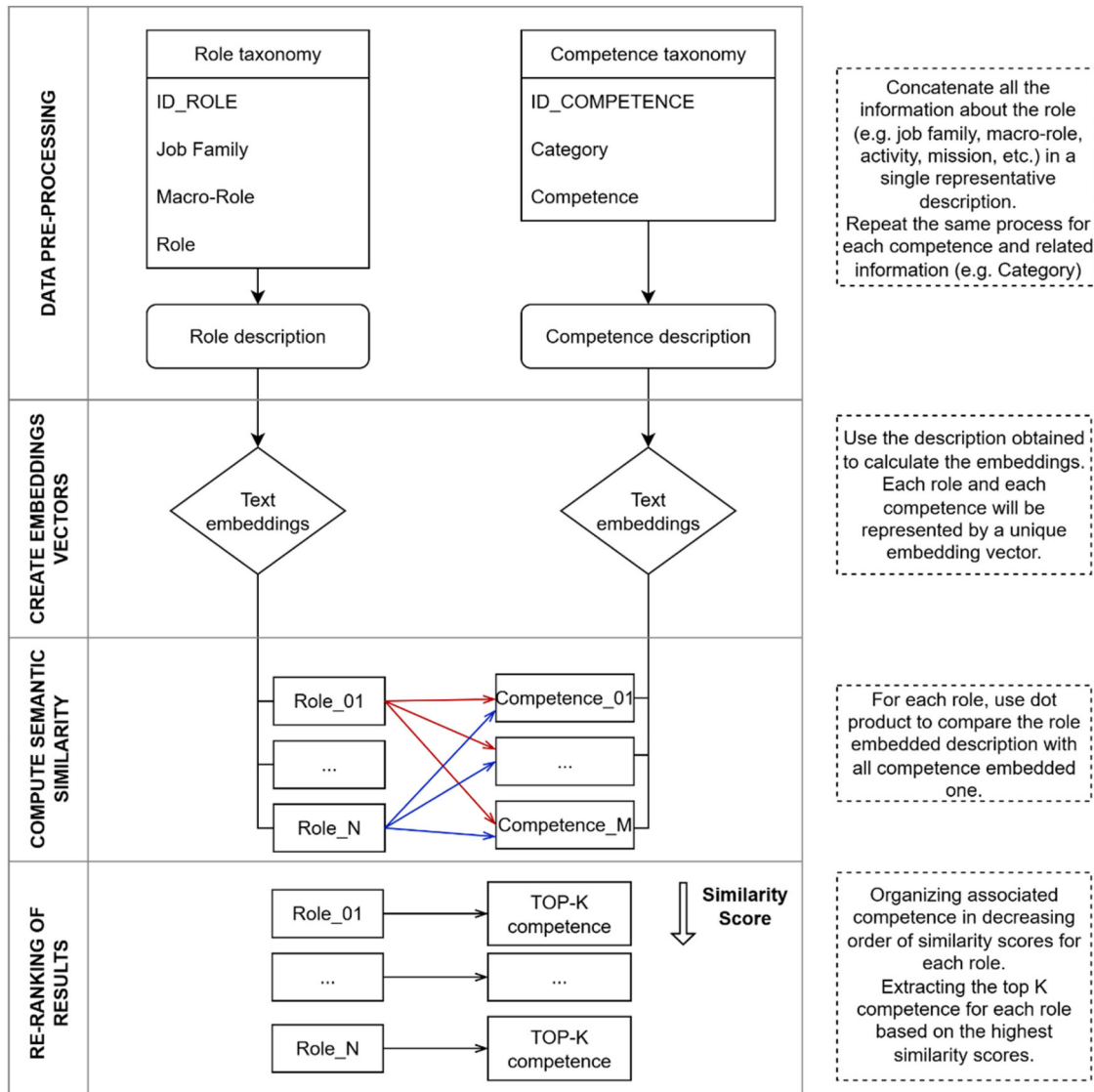


Fig. 4. RCE framework for role-competence association from corporate taxonomies.

semantic information and causing inefficiency in the process.

During the creation of the embedding vectors, the embedding models create separate vector representations for each role and competence concatenated description in an n-dimensional space that can range from 64 to 30522 dimensions, based on the selected model. This results in a set of representative vectors for each individual role and a set of vectors for each individual competence. This step is crucial for ensuring that each role and competence are represented accurately and consistently, thereby facilitating semantic comparison.

Subsequently, we calculated the semantic similarity using dot product, which is particularly effective in determining the similarity between the vectors within a multidimensional space (Kurek et al., 2024), between the role and competence embeddings vectors. In particular, it helps in determining the embedding of the first role of the professional role taxonomy with all the embeddings of the competence taxonomy. This process is repeated for each role within the role taxonomy.

Following the semantic similarity calculation, we performed the re-ranking step. For each role embedding, the final dataset was reorganised by placing the competence in the descending order of the similarity score. Lastly, we extracted only the top K competence for each role, based on the highest similarity scores. This process ensures that the competence associated with each role are the most relevant and appropriate ones, thereby improving the accuracy and efficiency of competence-based HRM.

Case study results

Industrial scenario description

The interviews conducted with the key informants of the selected company highlighted the challenges faced by the company and inform the development of the targeted solutions to enhance the integration and usability of role and competence taxonomies. In this section, we present the key outcomes emphasising the interplay between the insights derived from collaborative discussions and document analysis. There are two types of taxonomies: a competence taxonomy and a role taxonomy.

The role taxonomy comprises an initial division corresponding to the professional engineering families, a second hierarchical level corresponding to the macro-roles with descriptions, and the third level corresponding to the division of the roles, totalling 246 branches. Each role is accompanied by sufficiently comprehensive descriptions, including both strategic and operational aspects. Enrichment methodologies, such as ASTEM, are unnecessary for this specific area owing to the richness and level of detail of the role taxonomy. Although the role taxonomy is solid and well developed, it operates in isolation from the competence taxonomy.

The competence taxonomy comprises multiple ambiguous and overly sparse descriptions (see example in Table 4), making it ineffective for RCE. Therefore, the definitions for 766 competencies must be enriched using ASTEM. Previous studies and expert interviews demonstrated the complexity of obtaining precise and contextualised definitions for these competencies.

The company manually created a matrix associating each role with a set of up to 20 competencies; however, these taxonomies are subject to revision every two years, and the organizational requirements may change over time, making the manual process inefficient. This matrix supports strategic HRM by clearly defining the requirements across the organization. The matrix comprises a total of 3,954 records, with the

Table 4

Example of competence description that need semantic enrichment process.

Category	Electrification
Competence	Chemistry, Mechanical and electric systems
Competence Description	Management of mixed systems.

average number of competencies associated with roles categorised by the job family, as shown in Table 5. The matrix is constructed manually, which directly involves division managers and role members through a self-assessment activity presenting several significant limitations:

- The manual creation of the matrix requires considerable effort from managers, making the process lengthy and resource-intensive.
- The manual method is not sufficiently scalable, particularly for complex taxonomies or growing organizations.
- The matrix cannot be easily updated to reflect changes in the required competencies or role characteristics.

An analysis of the matrix construction process demonstrated the implementation of the RCE to significantly improve the connections between the roles and competencies.

Therefore, two core requirements of the company were identified through interviews and document analysis:

1. *Semantic enrichment of competence taxonomy*: A recurring theme was the requirement to enrich the sparse descriptions in a taxonomy to make it more actionable and to facilitate the integration between role taxonomy.
2. *Automated role-competence matrix*: The company highlighted the significance of automating the process of associating roles with competencies. This automation conserves save time, while enabling the company to dynamically update these mappings based on changes in the organizational requirements and taxonomy revisions.

Automating these processes reduces the manual effort and errors; furthermore, it enables the company to adapt rapidly to evolving demands in sectors with high level of complexity.

Frameworks Implementation

For the application of ASTEM and RCE in the case study, we employed advanced technological resources selected to ensure the effectiveness and scalability of the taxonomy enrichment and integration processes. In particular, we adopted technological solutions based on AI models and specific computational resources to satisfy the requirements of the various tests.

Competence taxonomy semantic enrichment using ASTEM

For the semantic enrichment of the competence taxonomy with ASTEM, we employed the GPT-4 model (OpenAI et al., 2023) based on the benchmark studies conducted by Chen et al. (2024) that highlight its advanced reasoning capabilities and ability to generate contextualised descriptions based on instructions, despite the absence of specific training (Excoffier et al., 2024). The GPT-4 model was particularly effective in producing coherent and contextualised definitions, thereby satisfying the specific requirements of the case study and overcoming the limitations of taxonomies lacking detailed textual content (OpenAI et al., 2023).

Therefore, we queried GPT-4 model (OpenAI et al., 2023) through

Table 5

Average (avg.) number of competencies associated with roles categorised by the job family.

Job Family	Avg. Competence per Role
Aeronautics Engineering (AE)	15
Electronics Engineering (EE)	17
Management Engineering (EM)	12
Mechanical Engineering (ME)	16
System Engineering (SE)	17
Software Engineering (SW)	17
Technology Management (TM)	18
Avg.	16

the OpenAI API to semantically enrich the competence taxonomy descriptions. In this case, GPT-4 generalises its knowledge to generate descriptions for competence within the taxonomy, despite the absence of explicit training on those specific elements. We developed a Python script to consider the competence taxonomy as input, extract the category, competence name, and original description from each taxonomy branch, and then query GPT-4 to semantically enrich the competence description based on the extracted labels. The documentation provided by OpenAI (OpenAI, 2023) helped in constructing a Python script to automate the API requests and adjust the process parameters, making the system efficient and robust. The system prompt specifies the task to be performed by the model and the method used to perform the task, whereas the user prompt provides only the hierarchical information corresponding to the competence description to be generated.

After generating the enriched competence descriptions, we randomly selected 10 outputs selected, which were subjected to an initial qualitative assessment by the partner company. This assessment utilised the control questions presented in Table 6, which were derived from the CRAAP test (Blakeslee, 2004; Newcastle University, 2024). However, the Currency (timeliness of data) and Authority (credibility of sources) criteria were omitted as they do not directly correspond to the evaluation of the quality of the content generated by GPT-4. This is because the validity of these aspects inherently corresponds to the training data of the model, which was considered sufficient to reference the documentation presented by OpenAI (OpenAI et al., 2023).

The remaining criteria, i.e., Relevance, Accuracy, and Purpose, were qualitatively assessed through a detailed review of each output from the randomly selected sample. This review was conducted during structured interviews, with key informants from the partner company involved in the case study.

Based on the feedback collected, the model prompt was iteratively refined until it reached its final version, as shown in Table 7.

A further analysis was conducted to determine whether the generated outputs (i.e., the descriptions of competencies) were consistent with the input data (i.e., the initial labels of the competence taxonomy).

The original study conducted by Excoffier et al. (2024) was applied in a hospital setting, where LLMs were used to generate textual datasets from ICD-10-CM code descriptions, a widely used taxonomy in U.S. hospitals; however, the core methodology remains valid in our context. The approach involves using generalist embedding models to verify whether the generated texts align with their corresponding classification codes.

In particular, the method relies on calculating the dot product similarity score using the E5-large-V2 embedding model, which was widely implemented for tasks requiring single-vector textual representations (Wang et al., 2022). In our case study, each ASTEM-generated competence description was compared against all the entries in the original competence taxonomy, and the competence identification code was assigned based on the highest similarity score. This comparison was performed to obtain the correct competence identification code for each generated description, thereby ensuring consistency between the input data and LLM-generated outputs. The same approach was applied to the

Table 6
Control questions extracted from the CRAAP test used to evaluate ASTEM output.

Criteria	Control Question
Relevance	Is the generated description at an appropriate level for the semantic enrichment based on the taxonomy labels task? Is the generated description useful because it provides a good definition?
Accuracy	Is the generated description reliable, truthful, or correct? Are there spelling or grammar mistakes in the generated description?
Purpose	Does the generated description sound objective and unbiased? Are there political, ideological, cultural, religious, institutional or personal biases in the generated description?

Table 7 System Prompt and User Prompt used in the ASTEM case study.	
System Prompt	Your task is to enrich a competence taxonomy by generating new competence descriptions. I will give you in input the catalogue structure: Competence_Category; Competence_name; Current_Competence_Description. Based on input information, write an impersonal short competence description (maximum 200 words). The description must NOT repeat Competence Name. If it possible, references to relevant engineering roles and technologies.
User Prompt	Competence_Category: {category} Competence_Name: {competence} Current_Competence_Description: {description}

original competence descriptions to determine the capability of embedding models to obtain the appropriate identification code, thereby enabling a comparison of the retrievability of ASTEM-generated descriptions versus the original ones.

The results of the semantic similarity analysis for validating the output presented in Table 8 indicated that, out of the 766 enriched competence descriptions, 73.4 % were associated with the correct identification code (Exact_Match), demonstrating the effectiveness of the ASTEM and GPT-4 in generating new coherent competence descriptions based on their labels. Additionally, 16.4 % of the enriched generated descriptions were associated with the wrong identification code, but fell within the correct competence category (Exact_Category). This indicates that there is some ambiguity in certain competencies within the taxonomy. In the remaining 10.2 % of cases, the association did not match despite the high values of the semantic similarity range (77 to 83).

We conducted a qualitative content analysis of two random competencies (i.e., radio communication and AFCS development for system engineering) to analyse the failure of No_match. The analysis indicated that the No_match enriched descriptions are primarily attributed to semantic ambiguity, loss of contextual specificity, and the limitations of the employed embedding model in differentiating the technical nuances. For instance, the enriched description for Radio Communication competence was incorrectly linked to Network Design, Planning & Engineering because they both share keywords such as 'radio' and 'TETRA' (see the complete descriptions in Table 9). Similarly, an enriched description corresponding to AFCS development for System Engineering was incorrectly associated with AFCS development for Aeronautic Engineering due to overlapping technical terms such as 'AFCS', despite the fact that the enriched descriptions comprised different aspects (see the complete descriptions in Table 10). However, this misalignment does not indicate that the enriched descriptions are off-topic or irrelevant. Instead, it highlights a limitation in the semantic analysis capabilities of the embedding model, which prioritises lexical and phrase similarity over precise conceptual alignment. This indicates that while the model can effectively identify typical thematic similarities, struggles with finer distinctions that require deeper contextual understanding.

Furthermore, we calculated the semantic similarity scores between the embeddings of each generated description and its original counterpart to assess whether the prompt constructed with ASTEM produced coherent outputs. The scores distribution presented in Fig. 5 tend to be high, indicating that the semantic enrichment process generated new descriptions very similar to the originals. This is likely because GPT-4, operating on a probabilistic basis, contained more information to draw upon, enabling it to produce coherent and consistent outputs. Conversely, the lowest similarity scores are associated with competencies that comprised very sparse original descriptions as demonstrated by the example depicted in Table 11. Consequently, GPT-4 generated outputs that varied more significantly from the originals although not necessarily producing incorrect content.

Table 8
Results of the output validation phase in the ASTEM case study.

Competence Taxonomy	Total Records	Exact_match	Exact_Category	No_match	Precision
Original	766	401	199	166	52.3 %
Enriched	766	562	126	78	73.4 %

Table 9
Comparison between radio communication and web communication and marketing competence descriptions.

Category	Electronics
Competence Name	Radio Communication
Original Competence Description	Knowledge of radio propagation theory, ability to use radio propagation tools, planning radio coverage system. Ability to verify and tuning radio communication system. Knowledge of 4 G and 5 G standards. Knowledge of DMR, TETRA and ATC standards (Radio aspects). Knowledge of Radio Frequency (RF) Communication-Electronics, including Radiotelegraphy (RTGY) and Radiotelephony (RTPY).

Enriched Competence Description

This competence involves a deep understanding of the principles and practical applications of electromagnetic wave transmission and reception. It encompasses the design, implementation, and optimization of systems that employ radio waves to exchange information over various distances without requiring physical connections. Proficiency in this area includes the ability to analyse and predict the behaviour of radio waves as they travel through different environments, which is crucial for ensuring reliable communication in various settings, from urban to rural and even in challenging terrains. Professionals with this competence can effectively utilise sophisticated software tools to model radio wave propagation and to create effective coverage maps that are essential for the planning and deployment of wireless networks. They can also effectively calibrate and adjust radio systems to satisfy the specific performance criteria, thereby ensuring compliance with international standards such as those governing 4 G and 5 G cellular networks, along with industry-specific protocols like digital mobile radio (DMR), terrestrial trunked radio (TETRA), and air traffic control (ATC) communication systems. In practice, individuals with this competence are essential for the fields of telecommunications, aerospace, defence, and public safety, where they contribute to the development and maintenance of critical communication infrastructures. Their expertise supports the seamless operation of everything from mobile phone networks to emergency responder radio systems and air traffic management communications.

Category	System Engineering
Competence Name	Network Design, Planning & Engineering
Original Competence Description	Knowledge of communication network, knowledge of network protocol, ability to design network architecture. Knowledge of 4 G and 5 G standards. Knowledge of DMR, TETRA, and ATC standards (Network aspects).

Enriched Competence Description

This competence requires the comprehensive understanding and application of principles to develop and implement efficient communication networks. It requires proficiency in various network protocols and the ability to create robust network architectures tailored to the specific organizational requirements. Individuals with this competence are expected to stay abreast of evolving technologies, including the intricacies of 4 G and 5 G standards, which are crucial for ensuring high-speed and reliable connectivity in modern telecommunications. Additionally, expertise in specialised communication standards such as DMR, TETRA, and ATC network requirements is essential. This knowledge is particularly relevant for ensuring interoperability and security in critical communication systems used by emergency services, public safety officials, and aviation authorities. Professionals in this field must be adept at planning, analysing, and optimising networks to satisfy the requirements of increased data traffic and the integration of new services. They must also be capable of conducting rigorous testing and validation to ensure that network performance meets both technical specifications and user expectations. The role often involves collaboration with cross-functional teams to align network design with business objectives, requiring strong communication and project management skills.

Role-competence association using RCE framework

The second step of the case study involves automating the linkage between roles and competence, thereby comparing the performance of two different embedding models: E5-large-V2 (Wang et al., 2022) and mxbai-embed-large-v1 (Xianming & Jing, 2023). The models were

Table 10
Comparison between AFCS development for system engineering and AFCS development for aeronautic engineering competence descriptions.

Category	System Engineering
Competence Name	AFCS development for System Engineering
Original Competence Description	Knowledge and application of methods and techniques for the Fly-By-Wire System design. Ability to design Automatic Flight Control Systems.

Enriched Competence Description

This competence requires a comprehensive understanding of the principles and practices associated with the design and implementation of automatic flight control systems (AFCS) within the realm of aerospace engineering. It requires proficiency in the application of advanced control theory, system dynamics, and real-time systems to develop reliable and efficient fly-by-wire systems that replace manual controls with an electronic interface. Professionals with this competence are adept at using specialised software and hardware tools to simulate and test AFCS designs before they are implemented in actual aircraft. They must also be knowledgeable of the stringent regulatory standards and certification processes that govern the safety and functionality of flight control systems. The role requires a collaborative approach, typically working closely with cross-functional engineering teams, including avionics, software, and hardware engineers, to ensure the seamless integration of the AFCS with other aircraft systems. A strong foundation in systems engineering principles is essential to oversee the project lifecycle from concept through to deployment and maintenance, ensuring that the AFCS meets all performance, safety, and reliability requirements.

Category	Aeronautics
Competence Name	AFCS development for Aeronautic Engineering
Original Competence Description	Knowledge and application of methods and techniques for the control laws and autopilot design. Development of design and write control laws functional requirements.

Enriched Competence Description

This competence requires a comprehensive understanding of the principles and methodologies used in the creation and implementation of AFCS. It requires the ability to conceptualise and define the functional requirements for control laws that govern the behaviour of the autopilot systems of an aircraft. Professionals with this competence can effectively design robust control strategies that ensure aircraft stability and performance, while also satisfying safety and regulatory standards. The role demands proficiency in systems engineering, along with a strong foundation in aerodynamics, control theory, and digital signal processing. Individuals must be capable of working with multidisciplinary teams, typically collaborating with software engineers, avionics specialists, and test pilots to translate theoretical control laws into practical, real-world applications. Familiarity with simulation tools and programming languages such as MATLAB, Simulink, or C++ is essential for modelling and validating the performance of the AFCS designs before they are implemented in hardware. As the field of aeronautics continues to evolve with new technologies, professionals in this area must stay abreast of the latest advancements in adaptive control, artificial intelligence, and machine learning to enhance the capabilities of autopilot systems and contribute to the development of more autonomous and efficient aircraft.

selected from the [Hugging Face repository](#) based on their ranking in the [Massive Text Embedding Benchmark \(MTEB\)](#), which evaluates models using several key metrics, including the model size, embedding dimensions, maximum token capacity, and memory usage (Muennighoff et al., 2023). These models were selected based on two main considerations: minimising computational costs (e.g., reducing memory usage and model size) while maximising the number of processable input tokens. The two pre-trained embedding models were selected to integrate the role and competence taxonomies within the RCE framework due to their strong semantic search capabilities and computational efficiency, even without domain-specific fine-tuning (see [Table 12](#) for details). Although both the models share similar strengths and limitations,

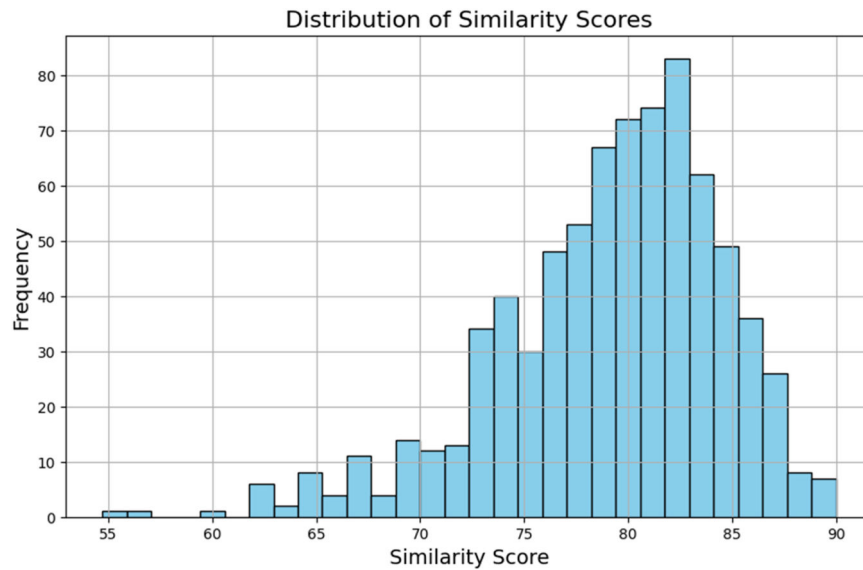


Fig. 5. Semantic similarity score between the enriched competence description and the original one.

E5-large-V2 has been extensively evaluated on widely recognised benchmarks, particularly for sentence similarity tasks (Excoffier et al., 2024). Conversely, mxlbai-embed-large-v1 has fewer publicly available test results, making it more difficult to evaluate its precise performance without direct experimentation. Consequently, this study aims to empirically compare their effectiveness in linking roles and competences.

These models were employed separately to vectorise the texts corresponding to each individual role within its taxonomy and for each competence in the respective taxonomy. The most appropriate competence sets for each role can be extrapolated by calculating the semantic similarity through the dot product between the representative vectors.

Following the RCE framework, the available data was pre-processed into the role taxonomy and the competence taxonomy, both in Excel format. All the information corresponded to a role, such as professional family, macro-role, and other textual information (mission, activities, and descriptions), was concatenated into a single cell to represent the role description (see an example in Fig. 6). A similar process was applied to each competence, creating a single text cell comprising descriptions and the category it belongs to (see example in Fig. 7). This approach enables the embedding models to better capture the overall semantic meaning, thereby improving the likelihood of accurate association.

The embedding models vectorised the role and competence descriptions in parallel. Subsequently, the representative vector of each role was compared with all the vectors representing the competence, thereby calculating the dot product score. The associated competencies were sorted in descending order of the similarity score for each role, and the top 16 competencies were extracted per role. These competencies were selected based on an average derived from the manually created role-competence matrix by the company (Table 5).

In this phase, we primarily aimed to test whether RCE combined with embedding models is a valid alternative for automatically creating role-competence matrix. The automatically generated role-competence pairs from two different embedding models (i.e. E5-large-V2 and mxlbai-embed-large-v1) was compared against the manually created ground truth (GT) dataset provided by the company.

The records are categorised into job families and the matching role-competence pairs are identified to evaluate the performance of the models (i.e. precision, recall, and F1-score), as shown in Table 13. This analysis was conducted on both the original competence descriptions and the ASTEM-enriched descriptions to determine whether semantic enrichment effectively enhanced the text, thereby making it more

semantically informative and reducing ambiguity in the embedding-based similarity analysis. The results indicate that semantic enrichment with ASTEM, combined with GPT-4, enhances the competence descriptions, making them more useful and interpretable for embedding models when compared with the original descriptions in the competence taxonomy. Furthermore, the comparative analysis between the two embedding models demonstrates that mxlbai-embed-large-v1 consistently outperforms E5-large-V2 in terms of the precision, recall, and F1-score across most job families. Notably, differences and patterns in the model performance metrics emerged across various job families. For instance, considering the job families, although the TM roles demonstrated higher precision in associating competencies, the EM roles exhibited lower precision, indicating that both models frequently suggested 'incorrect' competencies for EM roles despite high similarity scores. This discrepancy raised concerns regarding the potential biases, embedding model limitations, and structural differences between the descriptions of the roles.

To analyse this, we systematically extracted two EM and TM roles with the highest number of incorrect (false positive) competence associations to perform a qualitative comparative analysis. The number of common competencies between those selected automatically (RCE) and those defined manually (GT) is significantly lower in the EM roles, as shown in Table 14.

The EM roles (see an example in Table 15) present greater semantic complexity, as they include both strategic governance and operational management aspects. This caused ambiguity in the competence selection for the embedding models, where the models incorrectly prioritised operational and task-specific competencies (e.g., supplier management and work package definition) over broader engineering governance ones (e.g., system integration and project control). Conversely, the TM roles presented less ambiguity as they focused on innovation, technology scouting, and research collaboration. The embedding models were more successful in aligning competencies for these roles as their descriptions are clearer and less mixed with operational tasks. In fact, the embedding models are typically trained on datasets that contain rich documentation on technological innovation and business strategy (Excoffier et al., 2024; Wang et al., 2022), whereas they have less detailed coverage on specialised engineering knowledge (Pursnani et al., 2023; Roemer et al., 2024), thereby limiting the ability of the model to capture the EM role-specific competencies.

Another analysis was conducted to evaluate the total number of role-competence pairs identified by both the models and their alignment

Table 11

Example of comparison between original and enriched descriptions with low and high similarity score.

Category	Electrification
Competence Name	Chemistry, Mechanical and electric systems
Original Competence Description	Management of mixed systems.
Enriched Competence Description (Score: 54.7)	
This competence requires the comprehensive understanding and integration of chemical, mechanical, and electrical systems within various engineering applications. It requires a multidisciplinary approach to design, analyse, and manage systems where these three domains intersect. Professionals with this competence can effectively navigate the complexities of systems that rely on chemical processes, mechanical operations, and electrical controls. They can effectively ensure that the different components work harmoniously, optimising performance, and troubleshooting issues that arise from the interplay of these elements. In the context of electrification, this competence is particularly relevant for the development of batteries and fuel cells, where chemical reactions produce electrical energy, and mechanical components are essential for the structural integrity and thermal management of the system. It is also crucial in the design of electric vehicles and renewable energy systems, where mechanical and electrical engineering principles must be applied in tandem to achieve efficient energy conversion and utilization. Professionals with this competence are expected to stay abreast of technological advancements in materials science, electromechanical systems, and energy storage solutions. They typically collaborate with cross-functional teams, including chemical engineers, mechanical engineers, and electrical engineers, to innovate and improve system designs for enhanced sustainability and performance.	
Category	Materials & Structures
Competence Name	Additive Manufacturing
Original Competence Description	Knowledge and application of additive manufacturing approaches, including assessing the potential performance of additive approaches against a set of requirements. Knowledge and application of main techniques, tools and instrument of the additive manufacturing (e.g., 3D printings). Contribution to tasks adjacent to additive manufacturing, including broader technology assessment, defence anti-tamper, materials analysis, and test and component qualification.
Enriched Competence Description (Score: 82.5)	
This competence requires a comprehensive understanding of the processes and technologies used to create three-dimensional objects by adding material layer-by-layer, contrary to conventional subtractive manufacturing methodologies. Proficiency in this area includes the ability to evaluate and select appropriate materials, such as polymers, metals, and composites, for specific applications, along with the optimization of design for additive manufacturing (DfAM) principles to ensure efficient and effective production. Individuals with this competence are adept at operating and maintaining various additive manufacturing equipment, such as stereolithography (SLA), selective laser sintering (SLS), and fused deposition modelling (FDM) machines. They are also skilled in implementing computer-aided design (CAD) software to create digital models and prepare them for production. The role extends to the integration of additive manufacturing into the existing production workflows, thereby ensuring compliance with the industry standards and regulations. It is essential for roles such as materials engineers, design engineers, and manufacturing engineers involved in the aerospace, automotive, medical, and defence industries, where the ability to produce complex, lightweight, and custom parts can significantly enhance product performance and innovation.	

with GT, grouped by semantic similarity scores. Table 16 presents the results, which indicated that both mxbai-embed-large-v1 and E5-large-V2 exhibit similar performance patterns. In particular, the precision of role-competence associations improves significantly with the increase in the semantic similarity score.

E5-large-V2 does not present any matches in the lower ranges (60–75) due to its similarity scores clustering around higher values (0.7 to 1.0) because of the model architecture (Wang et al., 2022). Even if the scores are normalised (e.g., to a 0–1 scale), this process does not inherently make them comparable across models. The meaning of a normalised score can still vary based on the specific characteristics and training of the model. Therefore, for tasks such as text retrieval or semantic similarity, the relative ranking of the scores is more important than their absolute values (Wang et al., 2022).

The Venn diagram in Fig. 8 presents significant insights regarding

Table 12

Pre-trained embedding models used for RCE application in case study.

Model #1	E5-large-V2 (Wang et al., 2022)
Model #2	mxbai-embed-large-v1 (Xianming & Jing, 2023)
Description	Transformer-based sentence embedding models trained on multi-task data (e.g., natural language inference, question–answer retrieval, paraphrase detection, etc.) and designed for tasks involving semantic similarity, information retrieval, and other NLP applications where good dense text representations are necessary.
Key Features	<ul style="list-style-type: none">• Open-source models hosted on Hugging Face, providing transparent access to model weights and code under permissive licenses.• Their documentation is accessible and explains how to use, fine-tune, and evaluate them on various tasks.• They have relatively low computational requirements when compared with massive LLMs (e.g., GPT), making them accessible for most research or production settings.• Both are primarily fine-tuned for producing text embeddings (e.g., for semantic search or similarity). This narrow objective may constrain their ability to handle more general NLP tasks (e.g., detailed reasoning or context-heavy generation).• They have fewer parameters than the largest LLMs, which can reduce their ability to capture nuanced context or rare linguistic phenomena—especially on highly specialised or technical topics.
Limitation	<ul style="list-style-type: none">• Like most transformer-based encoders, they have a fixed token limit; text exceeding this limit must be truncated or split. This can cause loss of crucial context in longer documents.• They may underperform on niche scientific or domain-specific texts without further fine-tuning as they are general-purpose embedding models.

the performance and overlap of three sets (i.e., GT, E5-large-V2, and mxbai-embed-large-v1). The GT set contains 2397 unique elements, indicating a significant portion of data that neither model captures. Notably, the two models share 2527 elements, highlighting their effectiveness in capturing certain aspects of the role-competence data. Additionally, there are 1017 common elements across all three sets, indicating areas where both the models align well with the GT.

Discussion

Research implications

The results of this study help in establishing a knowledge base that can serve as a starting point for more in-depth studies in the future, making innovative AI technologies accessible and promoting informed use in the fields of competence-based HRM.

First, this study presents a detailed overview of the LLMs, including the fundamental principles of their architecture and functioning. Extensive research has been conducted (e.g., Vaswani et al., 2017; Wang et al., 2022) on explaining the working of these models; However, the previous studies were typically highly technical and limited research was conducted (e.g., Amin et al., 2022; Kurek et al., 2024) on applying competence-based HRM. Therefore, this study helps researchers and professionals without a specific background in computer science, as it elucidates complex concepts and presents a comprehensive understanding of the technologies underlying modern AI. By providing clear and in-depth explanations of the working of LLMs, the transformer architecture (Vaswani et al., 2017), and the role of embedding in capturing the semantic nuances of texts (Excoffier et al., 2024), this study helps in establishing a solid knowledge base that can facilitate further studies and applications in this field.

Furthermore, this study proposes ASTEM, which leverage the LLMs to generate enriched descriptions for corporate competence taxonomy elements, thereby emphasising the importance of prompt engineering in formulating requests to LLMs to obtain relevant and accurate responses (Chen et al., 2023). Although previous studies demonstrated that LLMs can generate structured text by capturing complex word relationships, they typically suffer from contextual inaccuracy and specificity issues

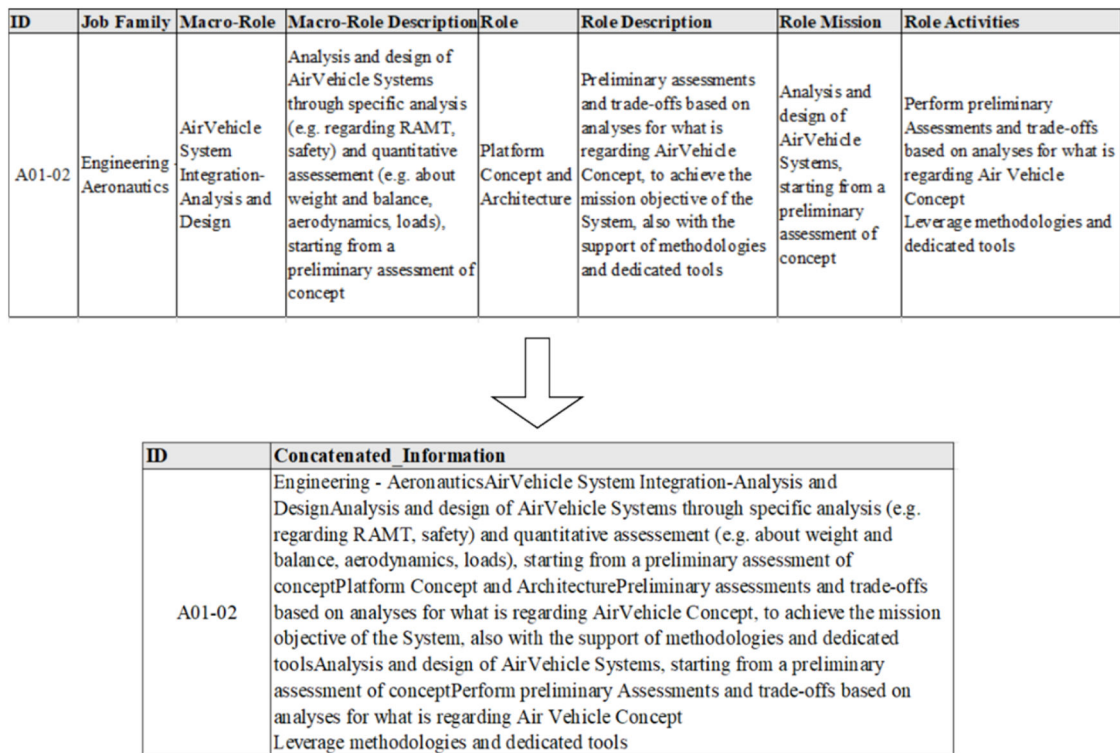


Fig. 6. Role taxonomy pre-processing for RCE implementation.

(Vartinen et al., 2022). ASTEM mitigates these limitations by integrating a preliminary taxonomy analysis, defining clear enrichment objectives to ensure that the LLM-generated descriptions align with the corporate taxonomies and prompt engineering techniques, thereby refining queries to optimise the LLM performance and improve the contextual accuracy (Chen et al., 2023; White et al., 2023). The iterative evaluation and prompt refinement, improves outputs to reduce hallucinations and factual inconsistencies (Lesage et al., 2024).

This methodology helps in reducing the manual effort required in taxonomy management and improving the contextual integration of various taxonomies (e.g., roles and competence). In particular, using LLMs to generate synthetic competence descriptions highlights a new application of generative AI in the context of HRM. The capability to generate accurate and contextualised enriched data can be extended to other types of taxonomy (e.g., technology, processes), thus presenting significant opportunities to improve the management recommendation systems even without the unavailability of the structured labelled dataset.

This study also proposes the RCE framework combined with embedding models to automate the association between the professional roles and required competence. The RCE reduces the time and resources required to create and maintain these associations, while also ensuring that the data remains up-to-date, regardless of changes to the company taxonomies. The RCE framework builds upon the previous studies conducted on embedding models for HRM (Amin et al., 2022; Kurek et al., 2024) but extends their application beyond simple job-description matching to create a fully automated taxonomy integration framework.

The automation of the role-competence association process represents a significant innovation in competence-based HRM, thereby providing a replicable model for future research. The comparison of E5-large-V2 (Wang et al., 2022) and mxbai-embed-large-v1 (Xianming & Jing, 2023) pre-trained embedding models demonstrates their differences in the semantic similarity tasks. This comparison serves as a benchmark for future studies aiming to evaluate the performance of other embedding models in role-competence association tasks and

similar applications.

The results of the RCE application with two different pre-trained embedding models also demonstrated that the ASTEM enriched competence taxonomy improves the effectiveness of the embedding models in creating accurate associations. This indicates that the embedding models can become fundamental tools to retrieve information with high accuracy if they are supported by adequate and rich textual content. Therefore, the effectiveness of the embedding model in capturing the semantic nuances between texts and creating relevant associations between the roles and competence could be a starting point to analyse other types of applications, such as the competence-technology association.

The ASTEM and RCE frameworks also contribute by extending the dynamic capabilities theory (Teece et al., 1997) and the KBV of the firm (Grant, 1996). Based on the dynamic capabilities perspective, the structured use of LLMs in HRM enhances the ability of an organization to sense, seize, and transform competence-related information. By improving the management of taxonomies and promoting a standardised competence-based approach, companies can dynamically adapt to evolving workforce requirements while avoiding dependence on biased or opaque decision-making processes. By automating the semantic enrichment of corporate competence taxonomies, ASTEM enables organizations to dynamically update and refine their competence datasets, thereby ensuring that newly emerging competencies and technologies are promptly incorporated. RCE further enhances the sensing and transformation capabilities of an organization by automating the identification and matching of competencies to roles, thereby improving the workforce adaptability and strategic alignment.

Additionally, from a KBV perspective, this study employs LLMs as a tool to enhance the ability of the firm to organise, structure, and utilise knowledge assets effectively. Furthermore, this study states that firms gain a competitive advantage through effective knowledge management (Grant, 1996). ASTEM employs KBV by structuring unstructured corporate knowledge into enriched competence descriptions, thereby making knowledge more accessible and actionable. The RCE framework

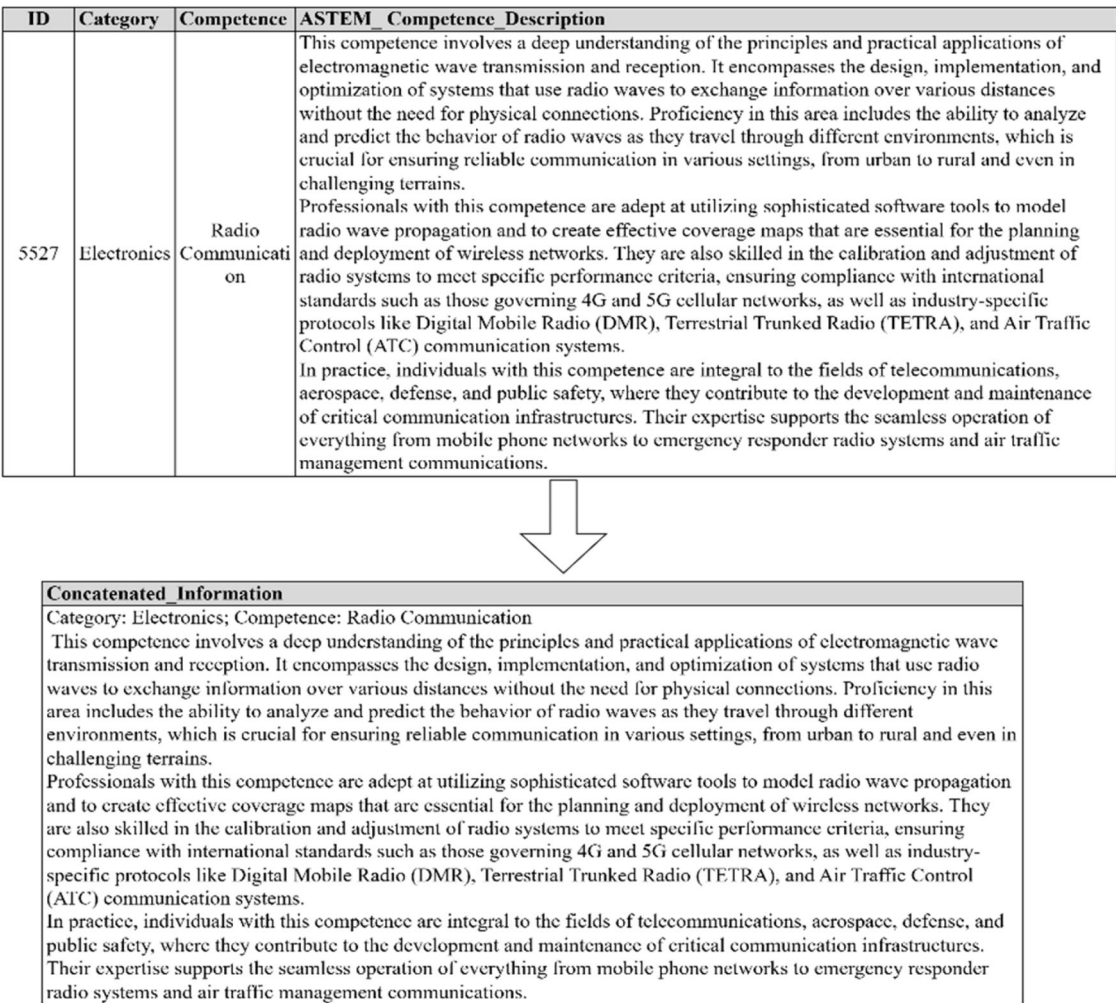


Fig. 7. Enriched competence pre-processing taxonomy for RCE implementation.

strengthens the KBV applications by employing AI-driven embeddings to systematically associate knowledge with the organizational roles, thereby reducing dependence on manual and the typically subjective knowledge mapping processes. Rather than delegating critical decision-making to AI-based models, the proposed approach ensures that these models enrich the corporate knowledge, thereby making information more accessible and actionable without compromising the ethical and legal considerations. By consciously limiting the applications of the LLMs to non-intrusive, but high-value knowledge structuring tasks, organizations can leverage the strengths of AI while mitigating the risks, thereby ensuring a balanced integration of human expertise and technological advancements.

Practical implications

In this study, we demonstrate how adopting LLMs can significantly transform the field of competence-based HRM, presenting new avenues for research and practical applications. In particular, the integration of advanced LLMs and prompt engineering represents a crucial step towards the automation and optimisation of competence management, promoting a more efficient usage of human resources and improving the business performance.

The implementation of the ASTEM and the RCE framework enables the automation of the semantic enrichment and association process between roles and competence enabling companies to significantly reduce the time and resources needed to keep their competence and roles taxonomies up to date, thereby improving operational efficiency

and information accuracy. These contributions enable the companies to save on human resources costs by reducing the need for manual interventions, which are typically time-consuming and resource-intensive. The conventional methods typically require subjective and manual evaluations, which can cause errors, inefficiencies and introduce subjectivity in the evaluation (Arslan & Cruz, 2022). Companies can achieve more accurate and consistent evaluations by leveraging LLMs, thereby eliminating many of the inefficiencies associated with the manual processes.

ASTEM includes iterative phases of evaluation and improvement, thereby ensuring that the generated descriptions and associations are constantly refined based on feedback. This continuous improvement process ensures that the adopted solutions remain relevant and effective over time.

In particular, companies can more precisely identify the ideal candidates for specific job positions and develop targeted training plans by using pre-trained embedding models to associate competence and roles through RCE. This process improves employee satisfaction as they are assigned to roles that best match their competence and aspirations. Additionally, targeted training plans help employees to further develop the competence required to excel in their roles, thereby increasing their motivation and reducing the turnover rates.

The ability to keep competence taxonomies updated enables companies to rapidly adapt to market changes and new technological requirements (Frierson et al., 2023). This is particularly relevant in high-tech sectors, where the evolution of the required competence is continuous and rapid. An updated and dynamic system enables the

Table 13
Performance metrics of role-competence pairing by job family.

E5-large-V2											
mxhai-embed-large-v1						Enriched Competence Descriptions					
Original Competence Descriptions						Enriched Competence Descriptions					
JF	GT	RCE	Match	Precision	Recall	F1-score	Match	Precision	Recall	F1-score	
TM	195	176	74	42.0 %	37.9 %	39.9 %	95	54.9 %	48.7 %	51.2 %	
AE	373	384	113	29.4 %	30.3 %	29.9 %	146	38.0 %	39.1 %	38.6 %	
EE	506	464	148	31.9 %	29.2 %	30.5 %	159	34.3 %	31.4 %	32.8 %	
SW	874	832	177	21.3 %	20.3 %	20.8 %	290	34.9 %	33.2 %	34.0 %	
ME	455	464	111	23.9 %	24.4 %	24.2 %	147	31.7 %	32.3 %	32.0 %	
EM	233	304	61	20.1 %	26.2 %	22.7 %	80	26.3 %	34.3 %	29.8 %	
SE	1318	1296	225	17.4 %	17.1 %	17.2 %	425	32.8 %	32.2 %	32.5 %	
	Total		909	Mean		26.6 %	Total	Mean		35.9 %	
				26.6 %		26.5 %		35.9 %		35.8 %	

Table 15

Example of RCE output for a EM-1 role from different embedding models and GT.

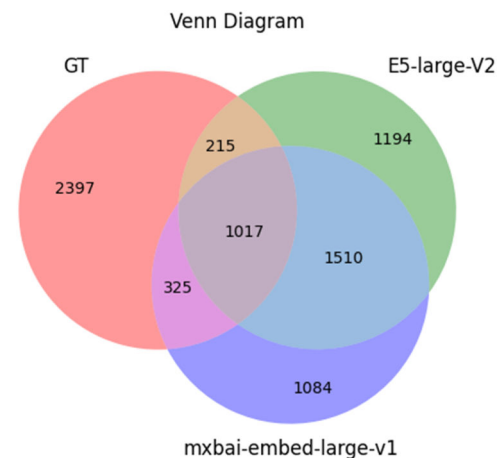
ID	EM-1
Role	Technical Disciplines Management and Activities Planning
Description:	
Governance of the overall product/project/program during the entire product/project/program lifecycle, ensuring the design authority integrity.	
In the frame of the IPT, delivery of the overall engineering solution through the management of the cross-disciplined engineering team, ensuring that the engineering team delivers an integrated solution compliant with the customer, contractual, legislative, and division requirements, also integrating all aspects related to performance, schedule, and cost perspective.	
<ul style="list-style-type: none"> • Organization of the WBS corresponding to a specific project/product, with consequent scheduling of all engineering activities and elaboration of periodical reports. • Definition of key performance indicators to measure the efficiency and effectiveness of engineering processes, with the elaboration of reports to support the process performance analysis. 	
Competence	R&T and Advanced Research Projects Management
E5-large-V2	This competence involves overseeing the lifecycle of the research and technology (R&T) initiatives, from conceptualization to completion. It requires the ability to orchestrate the preparatory phase, which includes defining the project scope, forming strategic partnerships, and negotiating terms within a consortium. The individual must possess the acumen to establish compelling project proposals that align with organizational goals and secure funding or support. Once a project is initiated, the focus shifts to the diligent management of its execution. This includes monitoring the progress against predefined targets corresponding to the cost, schedule, performance, and quality. The role demands a proactive approach to problem-solving and the capacity to adapt to evolving project requirements while aligning with the technology development objectives. The competence extends to ensuring that the outcomes of R&T projects are not only achieved but also effectively leveraged. This involves the preparation of an exploitation plan that defines the utilisation or commercialization of the results post-project. The individual must also ensure a smooth transition to project closure, capturing learnings and best practices for future endeavours. Professionals in engineering roles, such as R&D managers, project engineers, and technology strategists, typically exhibit this competence. They must be well-versed in project management methodologies and tools, as well as emerging technologies relevant to their field.
Competence	Business Performance Management and Process improvement
mxbai-embed-large-v1	This competence requires a comprehensive understanding of the intricate web of processes that support the engineering operations and their interdependencies with broader organizational functions. It requires the ability to critically analyse the current workflows, identify inefficiencies, and conceptualise enhancements that can lead to increased productivity and cost-effectiveness. Professionals who adept in this area are skilled in employing various methodologies, such as Lean, Six Sigma, and Continuous Improvement to drive process optimization. They are also proficient in setting and monitoring the performance metrics to ensure that the improvement efforts align with strategic business objectives. Mastery of this competence enables individuals to facilitate cross-functional collaboration, thereby fostering a culture of excellence and innovation that can significantly contribute to the competitive edge of the company. It is particularly relevant for roles such as Process Engineers, Quality Assurance Managers, and Operations Directors tasked with overseeing the seamless integration of engineering processes with the business's overall performance goals.
Competence	Concurrent engineering
GT	This competence requires the understanding and utilization of integrated approaches for product development, where cross-functional teams work collaboratively from the initial design phase through to production. It emphasises the parallelization of tasks to reduce the development time, improve the product quality, and reduce the costs. Professionals who adept in this area are skilled in coordinating various disciplines, such as design engineering, manufacturing engineering, and quality assurance, to ensure that product development considerations are addressed holistically. They are familiar with tools and methodologies that facilitate concurrent engineering, such as computer-aided design (CAD) and computer-aided manufacturing (CAM). Furthermore, they can leverage these technologies to optimise design for manufacturability and assembly (DFMA). By simultaneously considering all the elements of the product lifecycle, they can anticipate and mitigate potential issues in the early stages of design, thereby leading to more efficient processes and successful product launches. This competence is crucial in industries where the time-to-market and rapid innovation cycles are critical, such as automotive, aerospace, and consumer electronics.

Table 16

Comparison of embedding models based on the frequency of role-competence matches across different similarity score ranges.

Score Range	mxbai-embed-large-v1		E5-large-V2	
	Match	Frequency	Match	Frequency
60–65	13	14,13 %	0	0
65–70	138	28,57 %	0	0
70–75	571	30,05 %	0	0
75–80	479	37,99 %	236	21,99 %
80–85	120	65,22 %	969	34,29 %
85–95	11	68,75 %	27	72,97 %
Tot.	1342	34,10 %	1232	31,30 %

embedding models (specifically the open source E5-large-V2 and mxbai-embed-large-v1), supported by adequate semantic enrichment, can significantly improve the automatic association process between the roles and competence. This indicates that the quantity and quality of texts to be analysed are crucial for embedding models to perform well in semantic similarity analysis tasks. However, the overall precision of the two models indicates that there is still room for improvement, particularly in the lower score semantic similarity ranges. The possibility of reducing the human effort in validating the obtained results by establishing an optimal semantic similarity threshold further optimises the internal processes, enabling managers to focus on higher value-added

**Fig. 8.** Venn diagram with GT, E5-large-V, and mxbai-embed-large-v1 dataset.

activities. Identifying the threshold requires further research as embedding models are trained on different datasets and model textual information differently. Analysing other bigger models such as SFR-Embedding-Mistral (Meng et al., 2024), is an essential approach for achieving a more robust contribution on using embedding models to

Table 17
Comparative analysis of key aspects of HRM practices, ASTEM and RCE.

	Adaptability and Scalability	Bias & Subjectivity	Interpretability
Description	Updates in HRM information require considerable time and resources to adapt to different contexts, making it difficult to apply them across industries and as job roles evolve (Cao & Zhang, 2022)	HR decisions are typically subjective, influenced by individual experiences, cultural backgrounds, and unconscious biases (Kurek et al., 2024)	AI-based HRM tools typically function as 'black boxes', making recommendations difficult to justify or understand (Malik et al., 2023)
ASTEM			
Strengths	Reduces manual workload in taxonomy enrichment and updating with customised prompts tailored to specific domains	Competence descriptions are based on diverse datasets rather than individual opinions, promoting neutrality and standardization	Uses structured prompts and iterative validation to improve transparency and explainability
Limitations	Requires iterative refinements to ensure alignment with domain-specific contexts	May still inherit biases from the datasets used to train the LLMs	Still requires expert validation to ensure coherence and avoid misleading recommendations
RCE			
Strengths	Embedding capabilities are generalizable to different knowledge domains reducing manual work in role-competence matching	Embedding models automate role-competence matching, reducing human-driven inconsistencies and enhancing objectivity	Uses semantic embeddings to capture nuanced meanings in text, improving accuracy
Limitations	Effectiveness depends on the pre-trained model and the quality of input taxonomies	Biases in pre-trained embeddings may still affect role-competence associations	Interpretability remains an issue, as embeddings associations are based on semantic similarity of texts and not experience

create role-competence associations.

Although we performed some analyses on random samples regarding the bias of the GPT-4 for ASTEM using the CRAAP test (Blakeslee, 2004), the bias of the embedding models for RCE was not comprehensively analysed. Therefore, future works must analyse bias mitigation strategies, including explainable AI techniques to improve the transparency and interpretability (C.-S. Lin et al., 2023). Structured interviews with experts from the partner company help in evaluating the obtained ASTEM-enriched competence descriptions and RCE role-competence association from a more quantitative perspective (Taherdoost, 2022). Furthermore, an analysis based on clustering RCE and GT role-competence pairs is also performed to obtain a better understanding of this gap and the reasons behind the fact that the model does not fully reflect the GT dataset. However, increasing the trust in these models requires continuous collaboration between academia and industry to ensure their responsible development and usage. Implementing continuous evaluation mechanisms is essential for mitigating the risk of amplifying the existing societal biases and to promote fairness and equity in HR practices (Prabhakaran et al., 2022).

In summary, Table 18 presents the key future research directions to

Table 18
Future research directions and related research questions.

Future Research Directions	Research Questions
Generalizability of ASTEM and RCE Frameworks in Different Sectors	How can the ASTEM and RCE frameworks be adapted and validated in industries beyond aerospace, such as healthcare, technology, or finance?
Exploration of Larger Models for Role-Competence Association	How do larger models such as SFR-Embedding-2_R perform in automating role-competence associations when compared with the models used in this study? What is the optimal semantic similarity threshold for reducing human intervention?
Addressing Biases in LLMs	What are the common biases observed in LLMs when applied in competence-based HRM, and how can they be mitigated to ensure fairness and accuracy in role-competence matching?
Exploring Competence-Technology Associations	Can embedding models and semantic enrichment techniques applied to role-competence associations be extended to competence-technology associations?
Expanding the Application of LLMs in Competence-Based HRM	How can LLMs be applied to other aspects of HRM beyond role-competence matching, such as career path recommendation, competence forecasting, or performance assessment?
Understanding Model Interpretability in HRM Contexts	How can the interpretability of AI-driven role-competence association models be improved to provide more transparent and explainable outcomes for HRM professionals and decision-makers?

enhance the ASTEM and RCE frameworks corresponding to generalizability across various industries, the use of larger models, bias mitigation in LLMs, extensions to competence-technology associations, and improving AI interpretability for HRM applications.

Conclusions

In this study, we analysed the transformative potential of LLMs to improve competence-based HRM by conducting a specific case study in the industrial sector. We introduced two innovative solutions using LLMs: ASTEM and RCE frameworks. These solutions represent significant advancements in automating and enhancing competence-based HRM processes, thereby promoting more efficient and effective competence management.

This study contributes significantly to academic research in the fields of competence-based HRM and the application of LLMs in business contexts. In particular, we elucidated complex AI concepts and made them accessible to HRM professionals by providing a detailed overview of LLMs, including the principles of their architecture and functioning. The emphasis on prompt engineering is particularly notable, highlighting its significance in optimising the LLM performance to generate accurate and relevant responses. From a practical perspective, the proposed solutions help companies to use LLMs to semantically enrich the competence information based on their context of usage and to use embedding models to identify the set of competencies corresponding to a specific professional role. Companies can achieve greater operational efficiency, reduce costs, and improve employee satisfaction by automating the semantic enrichment of taxonomies and associating the roles with the corresponding competencies.

The findings of this study present a solid foundation for future analysis and broader applications of LLMs in HRM and beyond. The aerospace sector serves as a compelling case study owing to its complexity, dependence on high-level expertise, and extensive supply chain networks. However, the principles and methodologies presented in this study can be applied to other high-tech and knowledge-intensive industries. Similar challenges in competence management are observed

in other complex manufacturing sectors, where the integration of AI-driven competence frameworks could present comparable benefits. However, further research must be conducted to determine cross-industry applicability, identifying context-specific requirements and adapting LLM-based HRM solutions accordingly.

Future works must focus on optimising prompt engineering techniques, addressing biases in LLMs, and analysing larger and more diverse datasets to improve the robustness and accuracy of the role-competence associations. Additionally, comparative studies across different industrial domains could provide valuable insights on effectively adapting AI-driven competence management frameworks beyond the aerospace sector, extending their relevance to broader HRM and workforce planning strategies.

CRediT authorship contribution statement

Giuliana Barba: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Angelo Corallo:** Validation, Supervision, Methodology, Conceptualization. **Mariangela Lazoi:** Writing – review & editing, Validation, Supervision, Methodology, Conceptualization. **Marianna Lezzi:** Writing – review & editing, Validation, Supervision, Methodology, Investigation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The authors thank Massimo Scalvenzi and Chiara Spina for the valuable insights.

References

- Acha, V., Brusoni, S., & Prencipe, A. (2007). Exploring the miracle: Strategy and management of the knowledge base in the aeronautics industry. *International Journal of Innovation and Technology Management*, 04(01), 15–39. <https://doi.org/10.1142/S021987700700093X>
- Amin, M., Bergmann, J.-P., & Campbell, Y. (2022). Using vector representations for matching tasks to skills. In *Workshop on recommender systems for human resources 2022*. https://eur-ws.org/Vol-3218/RecSysHR2022-paper_11.pdf.
- Arsilan, M., & Cruz, C. (2022). Semantic enrichment of taxonomy for BI applications using multifaceted data sources through NLP techniques. In . 207 pp. 2424–2433). <https://doi.org/10.1016/j.procs.2022.09.533>
- August, T., Reinecke, K., & Smith, N. (2022). Generating scientific definitions with controllable complexity. In *Proceedings of the 60th annual meeting of the association for computational linguistics (Volume 1: Long Papers)* (pp. 8298–8317). <https://doi.org/10.18653/v1/2022.acl-long.569>
- Bafna, P., Shirwaikar, S., & Pramod, D. (2019). Task recommender system using semantic clustering to identify the right personnel. *VINE Journal of Information and Knowledge Management Systems*, 49(2), 181–199. <https://doi.org/10.1108/VJIKMS-08-2018-0068>
- Barba, G., Lazoi, M., & Lezzi, M. (2025). Enhancing technological taxonomies by large language models. In R. Zimmermann, J. C. Rodrigues, A. Simoes, & G. Dalmarco (Eds.), *Human-centred technology management for a sustainable future* (pp. 109–117). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-72494-7_12
- Baxter, P., & Jack, S. (2015). Qualitative case study methodology: Study design and implementation for novice researchers. *The Qualitative Report*. <https://doi.org/10.46743/2160-3715/2008.1573>
- Blakeslee, S. (2004). The CRAAP Test. *LOEX Quarterly*, 31(3). <https://commons.emich.edu/loexquarterly/vol31/iss3/4>
- Bousdekis, A., Lepenioti, K., Apostolou, D., & Mentzas, G. (2021). A review of data-driven decision-making methods for industry 4.0 maintenance applications. *Electronics*, 10(7), 828. <https://doi.org/10.3390/electronics10070828>
- Boyatzis, R. (1982). *The competent manager. A model for effective performance* (first edition, p. 1991). John Wiley & Sons. <https://doi.org/10.1002/smj.4250040413>
- Brin, D., Sorin, V., Vaid, A., Soroush, A., Glicksberg, B. S., Charney, A. W., Nadkarni, G., & Klang, E. (2023). Comparing ChatGPT and GPT-4 performance in USMLE soft skill assessments. *Scientific Reports*, 13(1), Article 16492. <https://doi.org/10.1038/s41598-023-43436-9>
- Brunello, G., & Wruuck, P. (2021). Skill shortages and skill mismatch: A review of the literature. *Journal of Economic Surveys*, 35(4), 1145–1167. <https://doi.org/10.1111/joes.12424>
- Brunton, S. L., Nathan Kutz, J., Manohar, K., Aravkin, A. Y., Morgansen, K., Klemisch, J., Goebel, N., Buttrick, J., Poskin, J., Blom-Schieber, A. W., Hogan, T., & McDonald, D. (2021). Data-driven Aerospace engineering: Reframing the industry with machine learning. *AIAA Journal*, 1–26. <https://doi.org/10.2514/1.J060131>
- Cao, C., & Zhang, Z. (2022). Machine learning-assisted competency modeling for Human resource management jobs. *Mobile Information Systems*, 2022, 1–15. <https://doi.org/10.1155/2022/8380307>
- Caratozzolo, P., Azofeifa, J. D., Mejia-Manzano, L. A., Rueda-Castro, V., Noguez, J., Magana, A. J., & Benes, B. (2023). A matrix taxonomy of knowledge, skills, and abilities (KSA) shaping 2030 labor market. In *2023 IEEE frontiers in education conference (FIE)* (pp. 1–8). <https://doi.org/10.1109/FIE58773.2023.10342955>
- Celikyilmaz, A., Clark, E., & Gao, J. (2021). *Evaluation of text generation: A survey* (arXiv: 2006.14799). ArXiv. <https://doi.org/10.48550/arXiv.2006.14799>
- Chang, E. Y. (2023). Prompting large language models with the Socratic method. In *IEEE Computing and communication workshop and conference (CCWC)*. <https://doi.org/10.48550/arXiv.2303.08769>
- Chen, B., Yi, F., & Varró, D. (2023). Prompting or fine-tuning? A comparative study of large language models for taxonomy construction. In *2023 ACM/IEEE International conference on model driven engineering languages and systems companion (MODELS-C)* (pp. 588–596). <https://doi.org/10.1109/MODELS-C59198.2023.00097>
- Chen, L.-F., & Chien, C.-F. (2011). Manufacturing intelligence for class prediction and rule generation to support human capital decisions for high-tech industries. *Flexible Services and Manufacturing Journal*, 23(3), 263–289. <https://doi.org/10.1007/s10696-010-9068-x>
- Chen, Y., Xu, B., Wang, Q., Liu, Y., & Mao, Z. (2024). *Benchmarking Large Language Models on Controllable Generation under Diversified Instructions* (arXiv:2401.00690). arXiv <https://arxiv.org/abs/2401.00690>
- Cilliers, P., & Spurrett, D. (1999). Complexity and post-modernism: Understanding complex systems. *South African Journal of Philosophy*, 18(2), 258–274. <https://doi.org/10.1080/02580136.1999.10878187>
- Clavié, B., & Soulié, G. (2023). *Large language models as batteries-included zero-shot ESCO skills matchers* (arXiv:2307.03539). arXiv. <https://doi.org/10.48550/arXiv.2307.03539>
- Corallo, A., Lazoi, M., Margherita, A., & Scalvenzi, M. (2010). Optimizing competence management processes: A case study in the aerospace industry. *Business Process Management Journal*, 16(2), 297–314. <https://doi.org/10.1108/14637151011035615>
- Corallo, A., Lazoi, M., & Secundo, G. (2012). Inter-organizational knowledge integration in Collaborative NPD projects: Evidence from the aerospace industry. *Knowledge Management Research & Practice*, 10(4), 354–367. <https://doi.org/10.1057/kmrp.2012.25>
- Creswell, J. W. (2022). *Research Design—Qualitative, quantitative, and mixed methods approaches (Sixth Edition)*. SAGE Publications. <https://uk.sagepub.com/en-gb/eur/research-design/book270550>
- Crowe, S., Creswell, K., Robertson, A., Huby, G., Avery, A., & Sheikh, A. (2011). The case study approach. *BMC Medical Research Methodology*, 11(1), 100. <https://doi.org/10.1186/1471-2288-11-100>
- Dahlkemper, M. N., Lahme, S. Z., & Klein, P. (2023). How do physics students evaluate artificial intelligence responses on comprehension questions? A study on the perceived scientific accuracy and linguistic quality of ChatGPT. *Physical Review Physics Education Research*, 19(1), Article 010142. <https://doi.org/10.1103/PhysRevPhysEducRes.19.010142>
- Dehbozorgi, N., & Parizi, R. M. (2023). Enhancing competency-based learning and employability through a course recommender AI-assisted system. In *The 24th annual conference on information technology education* (pp. 207–208). <https://doi.org/10.1145/3585059.3611435>
- Di Luozzo, S., D'orazio, L., & Schiraldi, M. M. (2021). In *Skills mismatch in operations & supply chain management roles: Perceptions from the European skills, competences, qualifications and occupations database*. September <https://www.scopus.com/record/display.uri?eid=2-s2.0-85124675416&origin=SingleRecordEmailAlert>
- Dickson, D. R., & Nusair, K. (2010). An HR perspective: The global hunt for talent in the digital age. *Worldwide Hospitality and Tourism Themes*, 2(1), 86–93. <https://doi.org/10.1108/17554211011012612>
- Dwivedi, S., Ghosh, S., & Dwivedi, S. (2023). Breaking the bias: Gender fairness in LLMs using prompt engineering and In-context learning. *Rupkatha Journal on Interdisciplinary Studies in Humanities*, 15(4). <https://doi.org/10.21659/rupkatha.v15n4.10>
- Eriksson, F., & Jönsson, A. (2023). *Automatic generation of definitions – Exploring if GPT is useful for defining words*. Linköping University-Department of Computer and Information Science.
- ESCO. (2017). *ESCO Strategic framework*. <https://ec.europa.eu/esco/portal/document/en/89a2ca9a-bc79-4b95-a33b-cf36ae1ac6db>
- Esposito, E., & Raffa, L. (2007). Global reorganisation in a high-technology industry: The aircraft industry. *International Journal of Globalisation and Small Business*, 2(2), 166. <https://doi.org/10.1504/IJGSB.2007.015480>
- Excoffier, J.-B., Roehr, T., Figueroa, A., Papaioannou, J.-M., Bressen, K., & Ortala, M. (2024). *Generalist embedding models are better at short-context clinical semantic search than specialized embedding models*. <https://doi.org/10.48550/arXiv.2401.01943> (arXiv:2401.01943). arXiv.
- Frierson, C., Wrobel, J., Senderek, R., & Stich, V. (2023). *Conceptualization of an AI-based skills forecasting model for small and medium-sized enterprises (SMEs)*. <https://doi.org/10.15488/13499>

- Furizal, Ramelan A., Adriyanto F., Maghfiro H., Ma'arif A., Kariyamin, Masitha A., & Fawait A.B. (2024). *Concerns of ethical and privacy in the rapid advancement of artificial intelligence: Directions, challenges, and solutions*. <https://doi.org/10.18196/jrc.v5i6.24090>.
- Goonawardene, N., Subashini, S., Boralessa, N., & Premaratne, L. (2010). A neural network based model for project risk and talent management. In L. Zhang, B.-L. Lu, & J. Kwok (Eds.), *Advances in neural networks—ISSN 2010* (pp. 532–539). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-13318-3_66. Vol. 6064.
- Grant, R. (1996). Toward a knowledge-based theory of the firm. *Strategic Management Journal*, 17, 109–122. <https://doi.org/10.1002/smj.4250171110>
- Gugani, A., & Misra, H. (2020). Implicit skills extraction using document embedding and its use in job recommendation. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(08), 13286–13293. <https://doi.org/10.1609/aaai.v34i08.7038>
- Gugani, A., Reddy Kasireddy, V. K., & Ponnalagu, K. (2018). *Generating unified candidate skill graph for career path recommendation* (pp. 328–333). <https://doi.org/10.1109/ICDMW.2018.00054>. 2018-NovemberScopus.
- Gupta, C. P., & Kumar, V. V. R. (2024). Artificial intelligence application in Human resource management: The way forward. In *2024 ASU International conference in emerging technologies for sustainability and intelligent systems (ICETISIS)* (pp. 1726–1730). <https://doi.org/10.1109/ICETISIS61505.2024.10459605>
- Iqbal, T., & Qureshi, S. (2022). The survey: Text generation models in deep learning. *Journal of King Saud University - Computer and Information Sciences*, 34(6), 2515–2528. <https://doi.org/10.1016/j.jksuci.2020.04.001>
- Jorzić, P., Yigit, A., Kanbach, D. K., Kraus, S., & Dabić, M. (2024). Artificial intelligence-enabled business model innovation: Competencies and roles of top management. *IEEE Transactions on Engineering Management*, 1–13. <https://doi.org/10.1109/TEEM.2023.3275643>
- Kekeya, J. (2021). *Qualitative case study research design: The commonalities and differences between collective, intrinsic and instrumental case studies* (pp. 28–37). DWU Press Publishers, 36.
- Kimble, C., De Vasconcelos, J. B., & Rocha, Á. (2016). Competence management in knowledge intensive organizations using consensual knowledge and ontologies. *Information Systems Frontiers*, 18(6), 1119–1130. <https://doi.org/10.1007/s10796-016-9627-0>
- Konrad, G., Khatwa, R., Feyereisen, T., & Kemp, A. (2022). Development of a flight deck human factors design and certification competency taxonomy: Preliminary findings. *Transportation Research Procedia*, 66(C), 279–291. <https://doi.org/10.1016/j.trpro.2022.12.028>
- Kurek, J., Latkowski, T., Bukowski, M., Świderski, B., Lepicki, M., Baranik, G., Nowak, B., Zakowicz, R., & Dobrakowski, L. (2024). Zero-shot recommendation AI models for efficient job–Candidate matching in recruitment process. *Applied Sciences*, 14(6), 2601. <https://doi.org/10.3390/app14062601>
- Lesage, J., Brennan, R., Eaton, S. E., Moya, B., McDermott, B., Wiens, J., & Herrero, K. (2024). Exploring natural language processing in mechanical engineering education: Implications for academic integrity. *International Journal of Mechanical Engineering Education*, 52(1), 88–105. <https://doi.org/10.1177/03064190231166665>
- Liévin, V., Hother, C. E., Motzfeldt, A. G., & Winther, O. (2023). *Can large language models reason about medical questions?* (arXiv:2207.08143). arXiv <http://arxiv.org/abs/2207.08143>.
- Lin, C.-S., Tsai, C.-N., Su, S.-T., Jwo, J.-S., Lee, C.-H., & Wang, X. (2023). Predictive prompts with joint training of large language models for explainable recommendation. *Mathematics*, 11(20), 4230. <https://doi.org/10.3390/math11204230>
- Lin, Y., Wang, X., & Xu, R. (2020). Semi-supervised human resource scheduling based on deep presentation in the cloud. *EURASIP Journal on Wireless Communications and Networking*, 2020(1), 73. <https://doi.org/10.1186/s13638-020-01677-6>
- Liu, P., Yuan, W., Fu, J., Jiang, Z., Hayashi, H., & Neubig, G. (2023). Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing. *ACM Computing Surveys*, 55(9), 1–35. <https://doi.org/10.1145/3560815>
- Madrid-García, A., Rosales-Rosado, Z., Freites-Núñez, D., Pérez-Sancristóbal, I., Pato-Cour, E., Plasencia-Rodríguez, C., Cabeza-Osorio, L., Abasolo-Alcázar, L., León-Mateos, L., Fernández-Gutiérrez, B., & Rodríguez-Rodríguez, L. (2023). Harnessing ChatGPT and GPT-4 for evaluating the rheumatology questions of the spanish access exam to specialized medical training. *Scientific Reports*, 13(1), Article 22129. <https://doi.org/10.1038/s41598-023-49483-6>
- Magron, A., Dai, A., Zhang, M., Montoriol, S., & Bosselut, A. (2024). *JOBSKAPE: A Framework for Generating Synthetic Job Postings to Enhance Skill Matching*. <https://doi.org/10.48550/arXiv.2402.03242> (arXiv:2402.03242). arXiv.
- Mahmoud, A. A., Al Shawabkeh, T., Salameh, W. A., & Al Amro, I. (2019). Performance predicting in hiring process and Performance appraisals using machine learning. In *2019 10th International conference on information and communication systems (ICICS)* (pp. 110–115). <https://doi.org/10.1109/ICICS.2019.8809154>
- Malik, A., Budhwar, P., & Kazmi, B. A. (2023). Artificial intelligence (AI)-assisted HRM: Towards an extended strategic framework. *Human Resource Management Review*, 33(1), Article 100940. <https://doi.org/10.1016/j.hrmr.2022.100940>
- Martínez, C. F., & Fernández, A. (2019). Ontologies and AI in recruiting. A rule-based approach to address ethical and legal auditing. In *Twenty-eighth international joint conference on artificial intelligence (IJCAI)*. <https://api.semanticscholar.org/CorpusID:211141929>.
- Mathew, V., Chacko, A. M., & Udhayakumar, A. (2018). Prediction of suitable human resource for replacement in skilled job positions using Supervised Machine Learning. In *2018 8th International symposium on embedded computing and system design (ISED)* (pp. 37–41). <https://doi.org/10.1109/ISED.2018.8704120>
- Mathis, R., Jackson, J., Valentine, S., & Meglitch, P. (2017). *Human resource management—15th*. <https://books.google.it/books?id=ihBCgAAQBAJ&print=c=frontcover&hl=it#v=onepage&q&f=false>.
- Meng, R., Liu, Y., Joty, S. R., Xiong, C., Zhou, Y., & Yavuz, S. (2024). SFR-embedding-mistral:enhance text retrieval with transfer learning. *Salesforce AI research Blog*. <https://blog.salesforceairesearch.com/sfr-embedded-mistral/>.
- Morandini, S., Fraboni, F., Hall, M., Quintana-Amate, S., & Pietrantonio, L. (2024). Human factors and emerging needs in aerospace manufacturing planning and scheduling. *Cognition, technology & work*. <https://doi.org/10.1007/s10111-024-00785-3>
- Muennighoff, N., Tazi, N., Magne, L., & Reimers, N. (2023). MTEB: Massive text embedding benchmark. In *EACL 2023 - 17th Conference of the European chapter of the association for computational linguistics, proceedings of the conference* (pp. 2006–2029). <https://doi.org/10.18653/v1/2023.eacl-main.148>
- Newcastle University. (2024). *Critical Evaluation Evaluating information generated by AI*. <https://www.ncl.ac.uk/mediav8/academic-skills-kit/file-downloads/Critical%20Evaluation%20of%20generative%20AI%20information.pdf>.
- Nguyen, M.-T., Nguyen, P.-T., Nguyen, V.-V., & Nguyen, Q.-M. (2021). Generating product description with generative Pre-trained Transformer 2. In *2021 6th International conference on innovative technology in intelligent system and industrial applications (CITISIA)* (pp. 1–7). <https://doi.org/10.1109/CITISIA53721.2021.9719940>
- Nikishina, I., Tikhomirov, M., Logacheva, V., Nazarov, Y., Panchenko, A., & Loukachevitch, N. (2022). Taxonomy enrichment with text and graph vector representations. *Semantic Web*, 13(3), 441–475. <https://doi.org/10.3233/SW-212955>
- OpenAI. (2023). *OpenAI guidelines*. <https://platform.openai.com/docs/api-reference/changelog>.
- OpenAI, Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F. L., Almeida, D., Altenschmidt, J., Altman, S., Anadkat, S., Avila, R., Babuschkin, I., Balaji, S., Balcom, V., Baltescu, P., Bao, H., Bavarian, M., Belgum, J., Zoph, B. (2023). *GPT-4 technical report* (arXiv:2303.08774). ArXiv. <https://doi.org/10.48550/arXiv.2303.08774>.
- Polverini, G., & Gregoric, B. (2024). How understanding large language models can inform the use of ChatGPT in physics education. *European Journal of Physics*, 45(2), Article 025701. <https://doi.org/10.1088/1361-6404/ad1420>
- Prabhakaran, V., Mitchell, M., Gebru, T., & Gabriel, I. (2022). *A Human rights-based approach to responsible AI* (arXiv:2210.02667). arXiv. <https://doi.org/10.48550/arXiv.2210.02667>.
- Pursnani, V., Sermet, Y., Kurt, M., & Demir, I. (2023). Performance of ChatGPT on the US fundamentals of engineering exam: Comprehensive assessment of proficiency and potential implications for professional environmental engineering practice. *Computers and Education: Artificial Intelligence*, 5, Article 100183. <https://doi.org/10.1016/j.caeai.2023.100183>
- Qin, C., Zhu, H., Xu, T., Zhu, C., Ma, C., Chen, E., & Xiong, H. (2020). An enhanced neural network approach to person-job fit in talent recruitment. *ACM Transactions on Information Systems*, 38(2), 1–33. <https://doi.org/10.1145/3376927>
- Ranade, N., Saravia, M., & Johri, A. (2024). Using rhetorical strategies to design prompts: A human-in-the-loop approach to make AI useful. *AI & SOCIETY*. <https://doi.org/10.1007/s00146-024-01905-3>
- Reagan, A. (2021). *Top challenges facing the future aerospace & defense workforce* (p. 5). <https://resources.gpstrategies.com/wp-content/uploads/2021/05/Article-TopChallengesFutureAerospaceWorkforce-05.25.21Olopt.pdf>.
- Roemer, G., Li, A., Mahmood, U., Dauer, L., & Bellamy, M. (2024). Artificial intelligence model GPT4 narrowly fails simulated radiological protection exam. *Journal of Radiological Protection*, 44(1), Article 013502. <https://doi.org/10.1088/1361-6498/ad1fd5>
- Roumeliotis, K. I., Tselikas, N. D., & Nasiopoulos, D. K. (2023). Unveiling sustainability in e-commerce: GPT-powered software for identifying sustainable product features. *Sustainability*, 15(15), Article 12015. <https://doi.org/10.3390/su151512015>
- Santana, M., & Díaz-Fernández, M. (2023). Competencies for the artificial intelligence age: Visualisation of the state of the art and future perspectives. *Review of Managerial Science*, 17(6), 1971–2004. <https://doi.org/10.1007/s11846-022-00613-w>
- Sheriff, N., & Sevukan, R. (2023). Discovering research data management trends from job advertisements using a text-mining approach. *Journal of Information Science*. <https://doi.org/10.1177/01655515231193845>. Scopus.
- Shivanjali, Singh, Dr. M., & Singh, Dr. T. (2019). Competency mapping: A strategic perspective in employee retention. *International Journal of Recent Technology and Engineering (IJRTE)*, 8(2), 3890–3898. <https://doi.org/10.35940/ijrte.A2196.078219>
- Skondras, P., Zotos, N., Lagios, D., Zervas, P., Giotopoulos, K. C., & Tzimas, G. (2023). Deep learning approaches for big data-driven metadata extraction in online job postings. *Information (Switzerland)*, 14(11). <https://doi.org/10.3390/info14110585>. Scopus.
- Stake, R. E. (1995). *The art of case study research (First)*. SAGE Publications Inc. https://books.google.it/books/about/The_Art_of_Case_Study_Research.html?id=ApGdX76b9kC&redir_esc=y.
- Taherdoost, H. (2022). *How to Conduct an Effective Interview; A Guide to Interview Design in Research Study Authors*. 11(1), 39–50. <https://hal.science/hal-03741838v1>.
- Tam, W. L., Liu, X., Ji, K., Xue, L., Zhang, X., Dong, Y., Liu, J., Hu, M., & Tang, J. (2022). *Parameter-efficient prompt tuning makes generalized and calibrated neural text retrievers*. <https://doi.org/10.48550/arXiv.2207.07087>.
- Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. *Strategic Management Journal*, 18(7), 509–533. [https://doi.org/10.1002/\(SICI\)1097-0266\(199708\)18<7509::AID-SMJ882>3.0.CO;2-Z](https://doi.org/10.1002/(SICI)1097-0266(199708)18<7509::AID-SMJ882>3.0.CO;2-Z).
- Turner, R. H. (2001). Role theory BT - handbook of sociological theory. *Handbook of sociological theory* (pp. 233–254).
- Ubeda, C. L., Santos, F. C. A., & Nagan, M. S. (2017). Analysis of the individual competences contributions to innovation management based on methodological

- triangulation. *Gestão & Produção*, 24(3), 595–609. <https://doi.org/10.1590/0104-530x1481-16>
- Vartinen, S., Hamalainen, P., & Guckelsberger, C. (2022). Generating role-playing game quests with GPT language models. *IEEE Transactions on Games*, 1–12. <https://doi.org/10.1109/TG.2022.3228480>
- Vasanthan, N., Shimran, M., Ahkam, A., Ishak, I., Silva, C., & Kuruppu, T. A. (2021). *Auto training an AI for detecting plant disease using twitter data annexed with a plant anthology* (pp. 437–442). <https://doi.org/10.1109/ICAC54203.2021.9671164>. Scopus.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). *Attention is all you need* (arXiv:1706.03762). arXiv. <https://doi.org/10.48550/arXiv.1706.03762>.
- Velásquez-Henao, J. D., Franco-Cardona, C. J., & Cadavid-Higuaita, L. (2023). Prompt Engineering: A methodology for optimizing interactions with AI-language Models in the field of engineering. *DYNA*, 90(230), 9–17. <https://doi.org/10.15446/dyna.v90n230.111700>
- Wallo, A., Kock, H., Lundqvist, D., & Coetzer, A. (2020). Understanding factors that enable and inhibit assessment of outcomes of competence development. *Human Resource Development Review*, 19(4), 384–421. <https://doi.org/10.1177/1534484320943332>
- Wang, L., Yang, N., Huang, X., Jiao, B., Yang, L., Jiang, D., Majumder, R., & Wei, F. (2022). *Text embeddings by weakly-supervised contrastive pre-training*. 1–17. <https://doi.org/10.48550/arXiv.2212.03533>.
- Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E., Le, Q., & Zhou, D. (2023). *Chain-of-thought prompting elicits reasoning in large language models* (arXiv: 2201.11903). arXiv. <https://doi.org/10.48550/arXiv.2201.11903>.
- White, J., Fu, Q., Hays, S., Sandborn, M., Olea, C., Gilbert, H., Elnashar, A., Spencer-Smith, J., & Schmidt, D. C. (2023). *A Prompt Pattern Catalog to Enhance Prompt Engineering with ChatGPT* (arXiv:2302.11382). arXiv <http://arxiv.org/abs/2302.11382>.
- Xianming, L., & Jing, L. (2023). AnglE-optimized text embeddings. In *Proceedings of ACL24*. <https://doi.org/10.48550/arXiv.2309.12871>
- Yin, R. K. (2017). *Case study research and applications: Design and methods (sixth edition)*. SAGE Publications.